

UNIVERSIDADE FEDERAL DO PARANÁ

LEONARDO RODRIGUES MINUCCI

REGIONALIZAÇÃO DE SÉRIE DE VAZÕES MENSAIS

UTILIZANDO REDES NEURAIS ARTIFICIAIS



CURITIBA

2014

LEONARDO RODRIGUES MINUCCI

REGIONALIZAÇÃO DE SÉRIE DE VAZÕES MENSAIS  
UTILIZANDO REDES NEURAS ARTIFICIAIS

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Engenharia de Recursos Hídricos, no Programa de Pós Graduação em Engenharia de Recursos Hídricos e Ambiental, Departamento de Hidráulica e Saneamento, Setor de Tecnologia, Universidade Federal do Paraná.

Orientador: Prof.Eloy Kaviski

CURITIBA  
2014

---

M668r

Minucci, Leonardo Rodrigues

Regionalização de série de vazões mensais utilizando redes neurais artificiais / Leonardo Rodrigues Minucci. – Curitiba, 2014.  
80f. : il. color. ; 30 cm.

Dissertação (mestrado) - Universidade Federal do Paraná, Setor de Tecnologia, Programa de Pós-graduação em Engenharia de Recursos Hídricos e Ambiental, 2014.

Orientador: Eloy Kaviski.

Bibliografia: p. 71-74.

1. Hidrologia - Paraná. 2. Redes neurais (Computação). 3. Análise multivariada. I. Universidade Federal do Paraná. II. Kaviski, Eloy. III. Título.

CDD: 551.48

---

## TERMO DE APROVAÇÃO

**LEONARDO RODRIGUES MINUCCI**

### **“Regionalização de série de vazões mensais utilizando Redes Naturais Artificiais”**

Dissertação aprovada como requisito parcial à obtenção do grau de Mestre, pelo Programa de Pós-Graduação em Engenharia de Recursos Hídricos e Ambiental do Setor de Tecnologia da Universidade Federal do Paraná, pela comissão formada pelos professores:

PRESIDENTE:



**Eloy Kaviski**  
Universidade Federal do Paraná  
Orientador

MEMBROS:



**Claudio Marchand Kruger**  
Universidade Positivo



**Miriam Rita Moro Mine**  
Universidade Federal do Paraná



**Mauricio Felga Gobbi**  
Universidade Federal do Paraná

**Curitiba, 21 de março de 2014**

“A diferença entre a inteligência e a estupidez é que a inteligência é limitada”  
Roberto Campos

## **AGRADECIMENTOS**

A Deus, por me fortalecer a cada dia.

À minha família, meus pais Minucci e Cecilia, e meu irmão Fredy, que sempre estiveram ao meu lado me apoiando.

À minha namorada Camila, que esteve comigo em todo o desenvolvimento desse trabalho.

Aos professores do PPGERHA, principalmente ao meu orientador Eloy Kaviski, pelo companheirismo e compreensão.

Ao meu amigo Gabriel, que em suas horas vagas me ajudou em detalhes que somente seus olhos percebem.

E a todos que, de alguma forma, contribuíram para a concepção dessa dissertação.

## SUMÁRIO

<b>LISTA DE TABELAS .....</b>	<b>VI</b>
<b>LISTA DE FIGURAS .....</b>	<b>VII</b>
<b>LISTA DE SIGLAS.....</b>	<b>IX</b>
<b>LISTA DE SÍMBOLOS.....</b>	<b>X</b>
<b>RESUMO.....</b>	<b>XII</b>
<b>ABSTRACT.....</b>	<b>XIII</b>
<b>1 INTRODUÇÃO .....</b>	<b>1</b>
1.1 CONSIDERAÇÕES GERAIS .....	1
1.2 MOTIVAÇÃO.....	3
1.3 OBJETIVOS DO ESTUDO .....	4
1.4 ORGANIZAÇÃO DO TRABALHO .....	4
<b>2 REVISÃO BIBLIOGRÁFICA.....</b>	<b>6</b>
2.1 RNAs E MODELAGEM CHUVA-VAZÃO .....	6
2.2 REGIONALIZAÇÃO hidrológica .....	7
<b>3 MÉTODO PROPOSTO .....</b>	<b>11</b>
3.1 CONSISTÊNCIA DOS DADOS FLUVIOMÉTRICOS .....	13
3.1.1 Curva-Chave.....	13
3.1.2 Série histórica de leitura de régua .....	19
3.2 MODELAGEM CHUVA-VAZÃO .....	21
3.2.1 Redes Neurais Artificiais (RNAs) .....	23
3.2.2 RNA do tipo <i>Multilayer Perceptron</i> – MLP .....	24
3.3 INTERPOLAÇÃO ESPACIAL.....	32
3.4 ANÁLISE ESTATÍSTICA MULTIVARIADA.....	34
3.4.1 Análise de componentes principais .....	35
3.4.2 Análise de agrupamentos .....	37
3.4.3 Análise discriminante .....	39
<b>4 APLICAÇÃO DO MÉTODO PROPOSTO AO ESTUDO DE CASO .....</b>	<b>41</b>
4.1 DISPONIBILIDADE DE DADOS.....	42
4.2 TRATAMENTO DOS DADOS .....	45
4.2.1 Climatologia .....	47
4.2.2 Fluviometria .....	47
4.2.3 Pluviometria .....	48
4.3 ESTRUTURAÇÃO DAS REDES NEURAS ARTIFICIAIS (RNAs) ..	48
4.4 VALIDAÇÃO DO MÉTODO DE REGIONALIZAÇÃO .....	51
<b>5 RESULTADOS .....</b>	<b>52</b>
5.1 CONSISTÊNCIA DOS DADOS FLUVIOMÉTRICOS .....	52
5.2 USO DA RNA PARA MODELAGEM CHUVA-VAZÃO .....	55
5.3 REGIONALIZAÇÃO HIDROLÓGICA.....	60
<b>6 CONCLUSÕES E RECOMENDAÇÕES .....</b>	<b>66</b>
<b>7 REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>71</b>
<b>APÊNDICE 1 – RESULTADOS DA MODELAGEM CHUVA-VAZÃO .....</b>	<b>75</b>

## LISTA DE TABELAS

TABELA 1: ENTRADAS PROPOSTAS NOS ESTUDOS DE MACHADO (2005) E MACHADO et. al. (2011) .....	6
TABELA 2: ESTAÇÕES FLUVIOMÉTRICAS E PLUVIOMÉTRICAS SELECIONADAS PARA O ESTUDO.....	43
TABELA 3: ESTAÇÕES PLUVIOMÉTRICAS UTILIZADAS NO ESTUDO.....	44
TABELA 4: ESTAÇÕES METEOROLÓGICAS UTILIZADAS NO ESTUDO. ....	47
TABELA 5: ENTRADAS PROPOSTAS PELO AUTOR (2014).....	50
TABELA 6: RESULTADOS DA ANÁLISE DE COMPONENTES PRINCIPAIS PARA UTILIZAÇÃO DAS ESTAÇÕES PLUVIOMÉTRICAS COMO ENTRADA PARA AS RNAS.....	56
TABELA 7: RESULTADOS OBTIDOS DA APLICAÇÃO DAS RNAS PARA MODELAGEM CHUVA-VAZÃO DAS BACIAS HIDROGRÁFICAS.....	59
TABELA 8: RESULTADO DA ANÁLISE DE COMPONENTES PRINCIPAIS DOS PARÂMETROS LEVANTADOS PARA AS ESTAÇÕES FLUVIOMÉTRICAS ESTUDADAS.....	60
TABELA 9: PARÂMETROS FÍSICOS E FISIOGRAFICOS OBTIDOS PARA AS ÁREAS DE DRENAGEM DAS ESTAÇÕES FLUVIOMÉTRICAS ANALISADAS. ....	62
TABELA 10: GRUPOS DE ESTAÇÕES FLUVIOMÉTRICAS FORMADOS.....	62
TABELA 11: RESULTADO DA ANÁLISE DISCRIMINANTE APLICADA ÀS BACIAS TESTE..	62
TABELA 12: DISTÂNCIAS E PRINCIPAIS ESTATÍSTICAS DA SÉRIE REGIONALIZADA PARA A ESTAÇÃO GUAMPARÁ. ....	63
TABELA 13: DISTÂNCIAS E PRINCIPAIS ESTATÍSTICAS DA SÉRIE REGIONALIZADA PARA A ESTAÇÃO MADEIREIRA GAVAZZONI. ....	64



## LISTA DE FIGURAS

FIGURA 1: DIAGRAMA DE FLUXO COM OS PASSOS PARA IMPLEMENTAÇÃO DO MÉTODO DE REGIONALIZAÇÃO DE REDES NEURAIS ARTIFICIAIS.....	12
FIGURA 2: CURVA-CHAVE DA ESTAÇÃO FLUVIOMÉTRICA RIO DA VÁRZEA, ESTADO DO PARANÁ.....	15
FIGURA 3: RELAÇÃO LINEAR ESTABELECIDADA PELO MÉTODO DE STEVENS PARA AS VAZÕES MAIS ALTAS MEDIDAS NA ESTAÇÃO FLUVIOMÉTRICA FAZENDA MARACANÃ, ESTADO DO PARANÁ. ....	18
FIGURA 4: ÁREA DE VALIDADE DA EXTRAPOLAÇÃO DE CURVA-CHAVE ESTABELECIDADA PELO MÉTODO DE STEVENS. ....	18
FIGURA 5: MUDANÇA DE RÉGUA OBSERVADA NA ESTAÇÃO ETA – GUARAPUAVA, ESTADO DO PARANÁ. ....	20
FIGURA 6: NEURÔNIO ARTIFICIAL.....	24
FIGURA 7: NEURÔNIO ARTIFICIAL.....	25
FIGURA 8: FUNÇÕES DE TRANSFERÊNCIA.....	31
FIGURA 9: MÉTODOS DE THIESSEN E DAS ISOIETAS (INVERSO DO QUADRADO DA DISTÂNCIA) APLICADO À BACIA HIDROGRÁFICA DO RIO PIRAPÓ, ESTADO DO PARANÁ. ....	34
FIGURA 10: VALORES DA SILHUETA PARA UM AGRUPAMENTO NÃO HIERÁRQUICO DE DADOS.....	39
FIGURA 11: ESTAÇÕES HIDROMETEOROLÓGICAS UTILIZADAS NO ESTUDO.....	46
FIGURA 12: CURVA DE DESCARGA OBTIDA PARA A ESTAÇÃO ETA – GUARAPUAVA EM SEU 3º PERÍODO DE RÉGUA (17/05/1995 A 31/12/2012). ....	53
FIGURA 13: CURVA DE DESCARGA OBTIDA PARA A ESTAÇÃO RIO DOS PATOS EM SEU 3º PERÍODO DE RÉGUA (01/11/2000 A 31/12/2012).....	53
FIGURA 14: HIDROGRAMA ANUAL DA ESTAÇÃO PONTE LEÔNCIO PRIMO.....	54
FIGURA 15: CURVA DE DESCARGA OBTIDA PARA A ESTAÇÃO PONTE LEÔNCIO PRIMO EM SEUS DOIS PERÍODOS DE RÉGUA DIFERENTES. ....	54
FIGURA 16: CALIBRAÇÃO E VALIDAÇÃO DAS RNAS PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO TEREZA CRISTINA. ....	58
FIGURA 17: CALIBRAÇÃO E VALIDAÇÃO DAS RNAS PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO FAZENDA MARACANÃ.....	58
FIGURA 18: GRÁFICO DA SILHUETA PARA TRÊS (ESQUERDA) E QUATRO (DIREITA) GRUPOS FORMADOS. ....	61
FIGURA 19: VAZÕES REGIONALIZADAS E OBSERVADAS – ESTAÇÃO GUAMPARÁ.....	65
FIGURA 20: VAZÕES REGIONALIZADAS E OBSERVADAS – ESTAÇÃO MADEIREIRA GAVAZZONI.....	65

FIGURA 1.1: CALIBRAÇÃO E VALIDAÇÃO DAS RNAS PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO PONTE DO GOIO BANG.....	75
FIGURA 1.2: CALIBRAÇÃO E VALIDAÇÃO DAS RNAS PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO PONTE LEÔNCIO PRIMO.....	76
FIGURA 1.3: CALIBRAÇÃO E VALIDAÇÃO DAS RNAS PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO ETA - GUARAPUAVA.....	76
FIGURA 1.4: CALIBRAÇÃO E VALIDAÇÃO DAS RNAS PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO BALSA DO CANTÚ.....	77
FIGURA 1.5: CALIBRAÇÃO E VALIDAÇÃO DAS RNAS PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO TORMENTA FAZENDA TRENTTO.....	77
FIGURA 1.6: CALIBRAÇÃO E VALIDAÇÃO DAS RNAS PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO PORTO CARRIEL.....	78
FIGURA 1.7: CALIBRAÇÃO E VALIDAÇÃO DAS RNAS PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO SALTO SAPUCAÍ.....	78
FIGURA 1.8: CALIBRAÇÃO E VALIDAÇÃO DAS RNAS PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO PORTO GUARANI.....	79
FIGURA 1.9: CALIBRAÇÃO E VALIDAÇÃO DAS RNAS PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO RIO DOS PATOS.....	79
FIGURA 1.10: CALIBRAÇÃO E VALIDAÇÃO DAS RNAS PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO PORTO SANTO ANTÔNIO.....	80
FIGURA 1.11: CALIBRAÇÃO E VALIDAÇÃO DAS RNAS PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO SÃO SEBASTIÃO.....	80

## LISTA DE SIGLAS

ACP: Análise de Componentes Principais

ADCP: *Acoustic Doppler Current Profiler*

ANA: Agência Nacional de Águas

COPEL: Companhia Paranaense de Energia

ETA: Estação de Tratamento de Água

HIDROWEB: *Website* da Agência Nacional das Águas

IPHMEN: Modelo Hidrológico do Instituto de Pesquisas Hidráulicas - UFRGS

MATLAB: *Mathematical Laboratory*

MLP: *Multi Layer Perceptron*

RMSE: *Root-Mean-Square Error*

RNA: Redes Neurais Artificiais

SIG: Sistema de Informação Geográfica

SMAP: *Soil Moisture Accounting Procedure*

## LISTA DE SÍMBOLOS

- $\mu$ : Parâmetro do algoritmo de otimização de Levenberg-Marquardt  
 $\beta$ : Parâmetro do algoritmo de otimização de Levenberg-Marquardt  
 $y$ : Vetor de saída da RNA  
 $x$ : Observações sujeitas à aplicação da Análise Discriminante  
 $g$ : Gradiente calculado para a aplicação do método de Levenberg-Marquardt  
 $J$ : Matriz jacobiana  
 $I$ : Matriz identidade  
 $H$ : Matriz de Hesse  
 $EW$ : Matriz de pesos atribuídos às entradas da RNA  
 $CW$ : Matriz de pesos atribuídos às camadas escondidas da RNA  
 $x$ : Valor de entrada a ser normalizado  
 $v$ : Velocidade do escoamento  
 $n$ : Número de meses analisados na modelagem chuva-vazão  
 $n$ : Coeficiente de rugosidade de Manning  
 $f(u)$ : Função de ativação da RNA  
 $e$ : Vetor de erro das RNAs  
 $d$ : Parâmetro de ajuste inserido na equação de Stevens para aumentar o grau de liberdade da equação  
 $a, a_1, a_2, a_3$ : Parâmetros de ajustes das curvas interpoladas  
 $T, T(t)$ : Temperatura e temperatura no mês corrente  
 $S$ : Declividade da linha de energia  
 $Q, Q(t), Q(t-1), Q(t-2)$ : Vazão, vazão nos meses corrente, um mês anterior e dois meses anteriores  
 $P(t), P(t-1), P(t-2)$ : Precipitação nos meses corrente, um mês anterior e dois meses anteriores  
 $EVT(t), EVT(t-1), EVT(t-2)$ : Evapotranspiração potencial nos meses corrente, um mês anterior e dois meses anteriores  
 $C$ : Coeficiente de Chezy  
 $A, A(h)$ : Área de escoamento da seção transversal ao curso d'água  
 $h$ : Cota  
 $\Sigma$ : Matriz de covariância  
 $\mu_k$ : Média aritmética dos elementos do grupo k formado na Análise Discriminante  
 $\pi_k, n_k$ : Número de elementos pertencentes a cada grupo formado na Análise Discriminante  
 $\lambda_1, \lambda_2, \dots, \lambda_k$ : Autovalores da matriz de entrada da Análise de Componentes Principais  
 $x_k$ : Variáveis da Análise de Componentes Principais  
 $w_1, w_2, \dots, w_n$ : Pesos atribuídos às entradas da RNA

$w_{k+1}, w_k$ : Variação no valor do peso da conexão na aplicação do método de Levenberg-Marquardt

$u_m$ : M-ésimo componente principal

$X, x_1, x_2, \dots, x_n$ : Dados de entrada da RNA

$x_{norm}$ : Dado de entrada normalizado para a RNA

$x_{min}$ : Valor mínimo dos dados de entrada da RNA

$x_{max}$ : Valor máximo dos dados de entrada da RNA

$w_0$ : Viés (*bias*) da RNA

$b_i$ : Distância mínima entre o i-ésimo indivíduo para os indivíduos dos diferentes grupos

$b^1, b^2, b^3$ : Viéses (*bias*) da RNA multicamadas

$a_i$ : Distância média entre o i-ésimo indivíduo aos outros indivíduos do mesmo grupo

$U_{rel}$ : Umidade relativa

$T_{min}$ : Temperatura mínima

$T_{max}$ : Temperatura máxima

$S_i$ : Valores de silhueta obtido para o j-ésimo indivíduo

$R_{solar}, R(t)$ : Radiação solar e radiação solar no mês corrente

$R_h$ : Raio hidráulico da seção transversal ao curso d'água

$R^2$ : Coeficiente de determinação

$Q_m$ : Vazão modelada a partir da RNA

$Q_0$ : Vazão observada

$L_k(x)$ : Função de verossimilhança do k-ésimo grupo

$G^k$ : Grupo k formado na Análise Discriminante

$D_k(x)$ : Distância de Mahalanobis

$h_0$ : Altitude mínima da régua limnimétrica instalada

## RESUMO

Pelo fato do Estado do Paraná ser uma das regiões mais propícias do país à instalação de aproveitamentos hidrelétricos – dada sua morfologia privilegiada e sua farta disponibilidade hídrica o ano todo – um dos pontos chave para o uso destes recursos é o conhecimento dos processos hidrológicos que nele ocorrem. Como não são todos os rios do estado que possuem estações de monitoramento fluviométrico, técnicas de regionalização hidrológica são geralmente empregadas para transferir informações para as regiões de interesse. Como muitas vezes tais técnicas envolvem abordagens subjetivas, dependentes da opinião do hidrólogo responsável, o objetivo desta dissertação foi propor um novo método de regionalização baseado em técnicas objetivas. O método consiste em regionalizar parâmetros de um modelo chuva-vazão, nesse caso Redes Neurais Artificiais (RNAs), para bacias hidrográficas carentes de medições fluviométricas, mas que apresentam medições históricas de chuva e temperatura. Para tornar objetivo o método, técnicas de análise estatística multivariada foram empregadas, como análise de componentes principais, análise de agrupamentos e análise discriminante. No trabalho foram estudadas 15 estações fluviométricas localizadas no interior do Estado e, destas, duas foram utilizadas como bacias hidrográficas teste para validação do método. Tanto os resultados da aplicação do modelo chuva-vazão como do método de regionalização foram bons, com índices de determinação e Nash-Sutcliffe elevados e erro médio quadrático baixo, quando confrontadas as vazões observadas e simuladas pelo modelo.

Palavras-chave: *hidrologia, regionalização, rede neural artificial, estatística multivariada.*

## ABSTRACT

Because of the state of Paraná is one of the most favorable regions of the country to the installation of hydroelectric plants - given its unique morphology and its plentiful water availability throughout the year - one of the key points for its development is the knowledge of its hydrological processes. As not all the rivers of the state have gaging monitoring, technical regionalization of hydrological stations are usually used to transfer information to the regions of interest. As very often such techniques involve subjective, dependent on the opinion of the hydrologist responsible, the goal of this thesis was to propose a new method for regionalization based on objective techniques. The method consists in regionalizing the parameters of a rainfall-runoff model, in this case Artificial Neural Networks (ANN), for needy catchments gauged measurements, but having historical measurements of precipitation and temperature. Thereby, multivariate statistical analysis techniques were performed, such as principal component analysis, cluster analysis and discriminant analysis. In this thesis, 15 gauged stations located within the state were studied, and among them, two basins were used as test method validation. Both the results of the application of rainfall-runoff model and the method of regionalization were satisfactory with high statistical indices when confronted flow rates observed and simulated by the model.

**Keywords:** *hydrology, regionalization, artificial neural network, multivariate statistics.*

# 1 INTRODUÇÃO

## 1.1 CONSIDERAÇÕES GERAIS

O Paraná é um Estado que merece destaque no que tange à disponibilidade hídrica e à morfologia apta à implantação de aproveitamentos hidrelétricos de pequeno porte, denominados de Pequenas Centrais Hidrelétricas, Centrais Hidrelétricas Geradoras e Micro Centrais Hidrelétricas.

Dentre os maiores desafios vivenciados em hidrologia atualmente, merece ênfase a dificuldade em conhecer adequadamente o comportamento dos processos hidrológicos característicos de determinadas regiões, principalmente devido à falta de dados. Muitas vezes, a avaliação de tais processos é realizada de maneira subjetiva e qualitativa, ou seja, muito dependente da opinião pessoal do hidrólogo e ligado apenas à natureza do processo. Nem sempre tal abordagem se faz eficaz, haja vista que sistemas ambientais funcionam de maneira complexa e quanto mais dados levantados em campo maior é a chance de se obter êxito no estudo.

É nesse contexto que se insere a regionalização hidrológica, ferramenta relativamente popular e utilizada para transferir informações hidrológicas de locais com dados para locais sem dados observados.

Por se tratar de um país com uma rede de monitoramento fluviométrico significativamente mal distribuída, pouco densa e de difícil operação, a aplicação de técnicas de regionalização hidrológica para previsão de vazão é frequentemente utilizada em estudos no setor hidroenergético brasileiro. Sua aplicação é vasta e vai desde estudos preliminares de verificação do potencial hídrico de uma região, até muitas vezes nas etapas de dimensionamento de obras e estruturas (Kaviski, 1992).

Dessa forma, métodos físicos e matemáticos, conceituais ou empíricos são desenvolvidos constantemente com a finalidade de contornar esse problema e suprir a carência de dados de vazão em bacias hidrográficas de maneira confiável. Segundo Tucci (2002), existem diferentes modelos que podem ser utilizados com essa finalidade.



Por serem mais bem distribuídas e com séries históricas geralmente de duração mais extensa, as redes de monitoramento pluviométrico e climático são muitas vezes utilizadas como entrada para a previsão de vazão em bacias carentes de postos fluviométricos, através de modelos matemáticos denominados chuva-vazão (Tucci, 1998).

Dentre as técnicas de modelagem chuva-vazão utilizadas, pode ser citada a Rede Neural Artificial – RNA, a qual vem sendo empregada na solução de vários problemas de recursos hídricos. RNAs são técnicas computacionais de inteligência artificial, as quais apresentam um modelo matemático não linear inspirado na estrutura neural de organismos inteligentes, que adquirem conhecimento através da experiência. Sua estruturação é feita a partir de elementos conhecidos por nós ou neurônios distribuídos espacialmente e que, interconectados, matematicamente representam uma função (Machado, 2005).

Machado (2005) aplicou RNA para modelagem chuva-vazão da bacia hidrográfica do rio Jangada, no Estado do Paraná. Diniz (2008) utilizou RNA para relacionar parâmetros hidrológicos obtidos de modelos chuva-vazão com características físicas e climáticas de bacias hidrográficas no nordeste brasileiro. Zhu *et. al.* (1994) utilizou RNAs em previsões de escoamentos de 1, 2, e 3 horas para a bacia do rio Butternut Creek em Nova York, Diniz e Clarke (2001) para regionalização dos parâmetros do modelo chuva-vazão SMAP mensal; Mazvimavi (2003) para a estimativa de características de vazões em bacias sem dados em Zimbabwe.

Além de RNAs, no processo de regionalização de vazão, outro importante ponto a ser examinado é a semelhança hidrológica e fisiográfica das regiões objeto de estudo, no caso bacias hidrográficas. É nesse cenário que se insere a análise estatística multivariada, sendo esta responsável por compreender, de maneira matemática, as similaridades e dissimilaridades entre os locais e, a partir disso, trazer resultados objetivos no tocante à viabilidade ou não da regionalização (Kaviski, 1992).

## 1.2 MOTIVAÇÃO

Em linhas gerais, no Estado do Paraná verifica-se que para as maiores bacias hidrográficas (rios Iguaçu, Ivaí, Tibagi e Piquiri) todo o potencial hidroenergético se encontra consolidado ou em fase final de estudos. Porém, muitas são as sub-bacias hidrográficas de pequeno e médio porte que ainda não foram investigadas e carecem de informações hidrológicas confiáveis à caracterização hidroenergética para fins de implantação de pequenos e médios aproveitamentos hidrelétricos.

Conforme Portaria nº 463 do Ministério de Minas e Energia (MME), publicada em 03 de dezembro de 2009, a qual dispõe sobre a metodologia para o cálculo dos montantes de garantia física de energia de usinas hidrelétricas não despachadas centralizadamente pelo Operador Nacional do Sistema Elétrico - ONS, para fins de participação no Mecanismo de Realocação de Energia – MRE, o histórico de vazões médias mensais das usinas hidrelétricas não devem ser inferior do que 30 anos e deve ser gerado de maneira que seja a mais extensa e atualizada possível.

Devido à indisponibilidade de dados de vazão em pequenas e médias bacias hidrográficas do Estado, somado com o fato da regulamentação pela Portaria MME nº 463 supracitada, métodos de regionalização hidrológica precisam ser desenvolvidos para a geração de série de vazões em estudos de centrais hidrelétricas.

Tanto na literatura nacional como internacional a grande maioria desses métodos para criação de série de vazões são baseados em técnicas subjetivas, as quais são geralmente baseadas em mapas e gráficos e dependem de elevado trabalho manual e de decisões subjetivas.

A proposta desse trabalho é a apresentação de um método objetivo para regionalização hidrológica de vazões médias mensais para pequenas e médias bacias hidrográficas baseado em técnicas de análise estatística multivariada e RNAs.

### 1.3 OBJETIVOS DO ESTUDO

O objetivo geral deste trabalho consiste em investigar a validade e significância da regionalização de RNAs treinadas e validadas para bacias hidrográficas as quais não dispõem de monitoramento fluviométrico, apenas pluviométrico e climatológico.

Os objetivos específicos são:

- Análise de consistência e tratamento dos dados pluviométricos e fluviométricos das estações hidrometeorológicas localizadas nas bacias hidrográficas analisadas;
- Calibração e validação da modelagem chuva-vazão a partir da variação de parâmetros de entrada das RNAs propostas para cada bacia;
- Regionalização de RNAs para bacias hidrográficas carentes de monitoramento fluviométrico a partir de análise estatística multivariada, comparando as características físicas (precipitação e temperatura) e fisiográficas (área de drenagem, declividade média, densidade de drenagem etc.) das bacias hidrográficas avaliadas nos itens anteriores.

### 1.4 ORGANIZAÇÃO DO TRABALHO

Essa dissertação é dividida em sete capítulos, sendo essa introdução o capítulo número 1.

No capítulo 2 apresenta-se a revisão bibliográfica, a qual faz um levantamento dos estudos encontrados na literatura sobre técnicas e métodos de regionalização hidrológica e treinamento e validação de RNAs para modelagem chuva-vazão.

No capítulo 3 apresenta-se o método proposto para regionalização hidrológica. Esse método faz uso da modelagem chuva-vazão a partir de RNAs e da análise estatística multivariada para verificação de padrões em bacias hidrográficas carentes de monitoramento fluviométrico no Estado do Paraná.

No capítulo 4 apresenta-se o estudo de caso em que o método proposto no capítulo 3 é aplicado e no capítulo 5 apresentam-se os resultados obtidos da aplicação do método.

No capítulo 6 são apresentadas as conclusões do trabalho e algumas recomendações para estudos futuros, e no capítulo 7 a bibliografia consultada ao longo do trabalho é apresentada.

## 2 REVISÃO BIBLIOGRÁFICA

A aplicabilidade de Redes Neurais Artificiais (RNAs) para modelagem chuva-vazão está praticamente consolidada no mundo atual, porém, não foram encontrados estudos que avaliam a possibilidade de utilização das RNAs em locais onde estas não foram treinadas nem validadas. Dessa forma, a revisão bibliográfica foi organizada em duas sub-seções que tratam dos temas RNAs e modelagem chuva-vazão e regionalização hidrológica.

### 2.1 RNAs E MODELAGEM CHUVA-VAZÃO

Machado (2005) e Machado *et. al.* (2011) fazem uma análise muito interessante ao aplicar dois diferentes métodos de modelagem chuva-vazão na bacia hidrográfica do rio Jangada, inserida no Estado do Paraná. Os autores confrontaram os resultados do modelo conceitual IPHMEN (Tucci, 1998), com o modelo empírico de RNAs para geração de séries sintéticas mensais de vazão e concluíram que o segundo obteve melhores resultados, implicando em maiores coeficientes de determinação e de Nash-Sutcliffe.

Diversos foram os parâmetros ajustados pelo autor nas RNAs a fim de melhorar o reconhecimento de padrões e obter melhores resultados. Foram variados o número de neurônios na camada intermediária, de reinicializações dos pesos e vieses (*epochs*), de dados de entrada e de validação, além de inicializados os pesos seis vezes para cada um dos modelos de entrada propostos, os quais estão apresentados na Tabela 1.

TABELA 1: ENTRADAS PROPOSTAS NOS ESTUDOS DE MACHADO (2005) E MACHADO *et. al.* (2011)

Modelo	Entradas	Saídas
1	$P(t).EVT^2(t)$	$Q(t)$
2	$P(t).EVT(t).Q(t-1)$	$Q(t)$
3	$P(t-1).P(t).EVT(t-1).EVT(t)$	$Q(t)$
4	$P(t-1).P(t).EVT(t-1).EVT(t).Q(t-1)$	$Q(t)$
5	$P(t-2).P(t-1).P(t).EVT(t-2).EVT(t-1).EVT(t)$	$Q(t)$
6	$P(t-2).P(t-1).P(t).EVT(t-2).EVT(t-1).EVT(t).Q(t-2).Q(t-1)$	$Q(t)$

P – PRECIPITAÇÃO MÉDIA MENSAL (mm), EVT – EVAPOTRANSPIRAÇÃO POTENCIAL MENSAL (mm) E Q – VAZÃO MÉDIA MENSAL (m<sup>3</sup>/s)

FONTE: MACHADO (2005); MACHADO *et. al.* (2011).

Nos estudos de Machado (2005) e Machado *et. al.* (2011) também é feita uma revisão bibliográfica a respeito do uso de RNAs para modelagem chuva-vazão em bacias hidrográficas distribuídas pelo Brasil e pelo mundo. Dentre os estudos citados pelo autor, destacam-se HSU *et.al.* (1995), Elshorbagy *et.al.* (2000), Anmala *et. al.* (2000), Tokar e Markus (2000), Sajikumar e Thandaveswara (1999), Lima e Ferreira Filho (2003) e Tokar e Johson (1999).

Destes autores, Anmala *et. al.* (2000), Tokar e Markus (2000), Lima e Ferreira Filho (2003) trabalharam com a previsão de vazão em escala mensal, a partir de dados como de precipitação, evapotranspiração e temperatura. Já HSU *et. al.* (1995) se destacaram por trabalhar com previsões em escala diária. Tanto nos estudos de HSU *et. al.* (1995) como de Tokar e Markus (2000) e Elshorbagy *et. al.* (2000) os resultados obtidos pelas RNAs foram confrontados com outros modelos e representaram melhores os eventos.

Maier e Dandy (2000) fazem uma abordagem completa de como as RNAs estão sendo cada vez mais utilizadas para prever não só vazão como também parâmetros hidrológicos em recursos hídricos. No artigo, todos os passos que devem ser seguidos para o desenvolvimento de tais modelos são descritos e avaliados a partir de uma revisão bibliográfica de 43 trabalhos.

Mais recentemente, Gomes e Montenegro (2010) apresentam uma análise dos dados pluviométricos e fluviométricos da área a montante do reservatório da UHE Três Marias, no rio São Francisco, bem como a calibração de um modelo de RNA para a previsão das vazões naturais afluentes em escala semanal. Dos resultados obtidos, verificou-se a eficiência no desempenho da RNA quando comparado com o modelo estocástico PREVIVAZ, desenvolvido pelo Centro de Pesquisa de Energia Elétrica (CEPEL) em 2004.

## 2.2 REGIONALIZAÇÃO HIDROLÓGICA

Kaviski (1992) classificou os métodos de regionalização em quatro grupos: (i) técnicas empíricas, (ii) técnicas de extensão de séries, (iii) técnicas de regressão e (iv) técnicas de mapeamento regional. Concluiu que as técnicas pertencentes aos diferentes grupos são complementares e as técnicas que fazem parte de cada grupo são competitivas.

No grupo das técnicas de mapeamento regional são incluídos os mapas de isolinhas, poligonais ou digitais e os métodos de metaregionalização, onde são incluídas as técnicas de estatística multivariada (análise de agrupamentos, análise discriminante, análise de componentes principais, análise de correlação canônica, etc.).

Mosley (1981) discutiu os problemas associados com a regionalização e verificou que a teoria e a prática da sua aplicação têm recebido a atenção de muitos pesquisadores na área científica. Das estratégias utilizadas para aplicação da regionalização, muitos pesquisadores têm definido sub-regiões homogêneas, utilizando resíduos de uma equação de regressão geral, como nos estudos de *Natural Environment Research Council* (1975) e Tasker (1982). Esses resíduos são plotados em um mapa da área em estudo e os limites são subjetivamente traçados de modo a criar grupos geograficamente contínuos e com valores semelhantes em sinal e valor. A subjetividade em tal abordagem significa que diferentes hidrólogos obteriam agrupamentos diferentes, ou que um conjunto de dados diferente (no tempo ou no espaço) teria resultado em diferentes regiões.

Mais recentemente, as técnicas estatísticas multivariadas, tais como análise de agrupamentos, têm sido utilizadas para identificar regiões hidrológicas homogêneas utilizando tanto as características hidrológicas como também características fisiográficas das bacias hidrográficas (Tasker, 1982; Acreman e Sinclair, 1986, Hughes, 1987, Haines *et. al*, 1988 *apud*. Mosley, 1981). Assumindo-se que a distribuição de probabilidade das características da bacia para cada região é normal multivariada, a uma bacia sem monitoramento pode ser atribuído um determinado agrupamento utilizando análise discriminante (Rao, 1965).

Geralmente, espera-se que os grupos formados compreendam bacias hidrográficas localizadas umas próximas geograficamente das outras, como verificado nos estudos de Mosley (1981) e Hughes (1987). Porém, nem sempre isso acontece, haja vista que os grupos são formados a partir de dados multidimensionais implicando, às vezes, em grupos sem significado geográfico algum.

Quando um método estatístico é selecionado para a determinação de regiões homogêneas no espaço, as variáveis explicativas das bacias hidro-

gráficas devem ter influência no fenômeno ou variável que se deseja regionalizar (Diniz, 2008). Muitas vezes, uma análise de componentes principais pode ajudar na seleção das variáveis, uma vez que a interpretação dos resultados obtidos nesse tipo de análise é geralmente feita sobre quantos poucos componentes principais podem explicar a variação total dos dados.

No Estado do Paraná, os estudos de regionalização elaborados no Centro de Hidráulica e Hidrologia Professor Parigot de Souza – CEHPAR conhecidos como HG-52, HG-77 e HG-171 merecem destaque (CEHPAR, 1989, 2006). Tais projetos foram desenvolvidos por interesse da Companhia Paranaense de Energia – COPEL em avaliar potenciais hidrelétricos de pequeno porte no Estado do Paraná.

Kaviski (1992) propôs um método de regionalização de curvas de permanência de vazões médias diárias para pequenas e médias bacias hidrográficas inseridas no Estado do Paraná. O método proposto fundamenta-se no seguinte: para todas as estações fluviométricas analisadas são ajustadas curvas de distribuição de probabilidade acumulada para a curva de permanência característica da estação. Com os parâmetros calculados, a partir de métodos estatísticos multivariados, como análise de agrupamentos e análise discriminante, verifica-se em qual grupo a bacia hidrográfica sem dados se insere e, a partir disso, determina-se sua curva de permanência pelos parâmetros regionalizados para a mesma.

A mesma ideia de análise de agrupamentos e análise discriminante foi aplicada em diferentes situações na região sul do país. Kaviski *et. al.* (1993) regionalizaram vazões médias no Estado de Santa Catarina e concluíram que os resultados obtidos eram limitados em função da amplitude das áreas de drenagem das estações com dados observados. Krüger e Kaviski (1994) ainda regionalizaram vazões máximas, médias e mínimas em bacias hidrográficas de até 5.000 km<sup>2</sup> inseridas no Estado do Paraná.

Cabe citar que métodos estatísticos não são os únicos utilizados para regionalização hidrológica; Diniz (2008) cita que existem ainda os métodos gráficos e os baseados em equações empíricas e em células quadradas, além dos métodos por delimitação geográfica, por mapeamento e da vazão índice.

Novaes *et. al.* (2007) avaliaram o desempenho de cinco métodos de regionalização de séries históricas de vazão na bacia hidrográfica do rio Para-



catu, em Minas Gerais. As técnicas selecionadas foram: método baseado na interpolação linear, método proposto por Chaves *et. al.* (2002) *apud.* Novaes *et. al.* (2007), método da interpolação linear modificado e método Chaves modificado. Os quatro primeiros métodos são baseados única e exclusivamente nas relações existentes entre as áreas de drenagem dos dois postos selecionados, já os dois últimos levam ainda em consideração valores de precipitação média anual na área de drenagem das seções de interesse. Dos resultados obtidos, verificou-se que as vazões estimadas pelos cinco métodos de regionalização não apresentaram diferenças expressivas nos resultados obtidos e que a inserção da variável precipitação nos dois últimos métodos não apresentou melhorias que justificassem a sua recomendação.

Dessa maneira, por conta dos bons resultados obtidos nos últimos anos ao utilizar o modelo de RNAs para previsão de vazão, somado ao fato de terem sido gerados resultados estatisticamente satisfatórios quando testados em bacias do Estado do Paraná, optou-se pela utilização deste modelo para a previsão de série de vazão a partir de dados climatológicos e pluviométricos.

Com respeito às técnicas de estatística multivariada para estabelecer as regionalizações, a escolha da análise de agrupamentos seguida de análise discriminante foi feita por conta dos bons resultados em trabalhos realizados por Kaviski (1992), Krüger e Kaviski (1994) e CEHPAR (1989,2006) nos últimos anos no Estado do Paraná.

### 3 MÉTODO PROPOSTO

O método proposto tem por objetivo a regionalização de Redes Neurais Artificiais (RNAs) para bacias hidrográficas de pequeno porte, com áreas de drenagem situadas entre 100 km<sup>2</sup> e 5.000 km<sup>2</sup>.

Esse método permite que para bacias hidrográficas carentes de monitoramento fluviométrico sejam obtidas séries de vazões médias mensais a partir de RNAs calibradas e validadas em bacias hidrográficas com características fisiográficas e climatológicas estatisticamente semelhantes.

A estrutura do método é baseada em técnicas de análise e consistência de dados de medição hidrometeorológica, modelagem chuva-vazão a partir de RNAs, análise de componentes principais, análise de agrupamentos e análise discriminante.

Inicialmente são obtidas as características fisiográficas das bacias hidrográficas com monitoramento fluviométrico, como área de drenagem em km<sup>2</sup>, centróide em coordenadas UTM, altitude da estação em metros, declividade média do talvegue principal em m/km e densidade de drenagem em km/km<sup>2</sup>. Após isso, são realizadas análises de consistência dos dados de precipitação das estações pluviométricas inseridas na bacia hidrográfica e dos dados de vazão do posto fluviométrico, de modo a obter uma série histórica mensal de precipitação e vazão e de parâmetros explicativos da bacia, como a precipitação média anual de longo período na bacia, em mm, e seu coeficiente de variação. Complementarmente, através do método de interpolação pelo inverso do quadrado da distância, são obtidas para o centróide da bacia hidrográfica séries históricas mensais de medições climatológicas, como dados de temperatura, temperatura máxima, temperatura mínima, umidade relativa e radiação solar, os quais serão utilizados na etapa seguinte, de modelagem chuva-vazão.

Para as bacias hidrográficas são ajustados modelos de RNAs que simulam o comportamento hidrológico da bacia de maneira estatisticamente satisfatória, verificada por coeficientes comuns na literatura. Como entradas são utilizados os dados de precipitação e temperatura obtidos na etapa anterior, os quais tentam prever as vazões, também obtidas na etapa anterior.

Uma vez definida a melhor RNA para cada uma das bacias hidrográficas analisadas, tem início a técnica de análise estatística multivariada com vistas à regionalização das RNAs a partir de dados físicos e fisiográficos. Adicionando na amostra de observações as características da bacia hidrográfica que não dispõe de dados de vazão, é realizado um agrupamento não hierárquico de modo a verificar em qual região homogênea a mesma se enquadra. Finalmente, realiza-se uma análise discriminante para seleção da RNA a ser utilizada na simulação chuva-vazão dessa bacia.

Na Figura 1 é apresentado um diagrama onde está esquematizada a sequência de técnicas utilizadas e definidas para o método proposto.

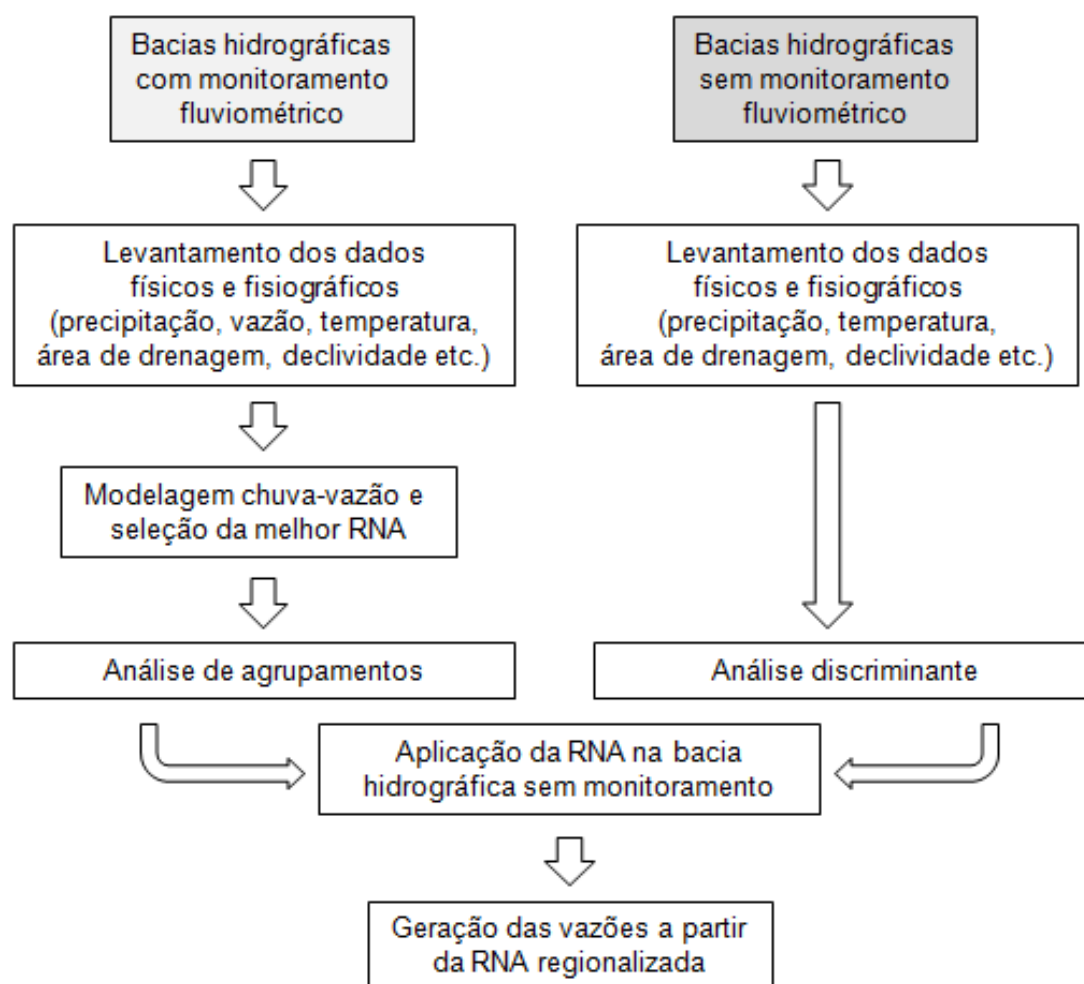


FIGURA 1: DIAGRAMA DE FLUXO COM OS PASSOS PARA IMPLEMENTAÇÃO DO MÉTODO DE REGIONALIZAÇÃO DE REDES NEURAIS ARTIFICIAIS.

FONTE: O AUTOR (2014).

Este capítulo está subdividido em quatro itens, que apresentam as técnicas e o embasamento teórico dos métodos propostos para esse trabalho, são eles: consistência de dados fluviométricos, modelagem chuva-vazão, mapeamento regional e análise estatística multivariada.

### 3.1 CONSISTÊNCIA DOS DADOS FLUVIOMÉTRICOS

Para Sefione (2002), no ramo da ciência, um dos aspectos mais importantes a ser tratado é a confiabilidade dos dados com os quais serão elaboradas novas ideias e hipóteses sobre os mais diversos temas. A compreensão dos mecanismos de ação de um fenômeno físico está diretamente relacionada com o processo de aquisição de dados para a explicação do próprio fenômeno, quanto menor for o grau de incerteza do dado concebido, maior a chance de se obter êxito e precisão na análise.

No ramo da hidrometria - ciência que mede e analisa características físicas e químicas da água através de métodos, técnicas e instrumentação utilizados em hidrologia - a busca de maior precisão nas medições realizadas tem uma importância fundamental no sentido de minimizar incertezas.

Qualquer que seja a finalidade de um estudo de disponibilidade hídrica que faça uso de dados de vazão em cursos d'água, uma ferramenta de extrema importância que deve ser cuidadosamente analisada para o trecho do rio é a curva-chave.

#### 3.1.1 Curva-Chave

Caracterizada por representar graficamente a relação que existe entre a descarga medida e a leitura simultânea de cota em uma determinada seção transversal, a curva-chave é uma função que envolve características geométricas e hidráulicas do trecho de um curso d'água.

A partir de uma série de dados de descarga levantados em campo para uma determinada seção transversal, juntamente com processos iterativos de minimização de erros, torna-se possível a obtenção de uma expressão matemática para curva-chave.

As hipóteses por trás da validade desta suposição de relação biunívoca são as de que em qualquer instante, na seção transversal do trecho do rio em questão, o escoamento é aproximadamente permanente (invariante no tempo) e uniforme (acelerações desprezíveis). Tais hipóteses só são verdadeiras se a taxa de variação do nível do rio com o tempo é relativamente pequena, e a seção transversal do rio, próxima à estação de medição, varia lentamente ao longo do eixo do rio.

Sempre que uma curva-chave é válida, existe uma relação entre várias características da geometria da seção transversal do canal onde se encontra o posto de medição. As duas equações matemáticas mais conhecidas que propõem esta relação são as equações de Manning [Equação (1)] e Chèzy [Equação (2)]:

$$v = \frac{R_h^{\frac{2}{3}} \cdot S^{\frac{1}{2}}}{n} \quad (1)$$

$$v = C \sqrt{R_h \cdot S} \quad (2)$$

onde  $v$  é a velocidade do escoamento (m/s),  $R_h$  é o raio hidráulico da seção (m), definido pela relação entre a área e o perímetro molhado,  $S$  é a declividade da linha energia (m/m),  $n$  é o coeficiente de rugosidade de Manning ( $s \cdot m^{-1/3}$ ) e  $C$  é o coeficiente de Chezy ( $m^{1/2}/s$ ).

Segundo Maidment (1992), o conhecimento em hidráulica teórica contribui para a determinação da forma geral de uma curva-chave. Em canais longos e retilíneos, onde a vazão é controlada apenas pelo atrito da superfície de contato com o fluido, a curva-chave segue uma curva de potência baseada na seguinte equação:

$$Q = a(h - h_0)^b \quad (3)$$

onde  $Q$  é a vazão líquida ( $m^3/s$ ),  $a$  e  $b$  são coeficientes de ajuste da equação e  $h$  e  $h_0$  são as leituras de régua (cm), correspondentes às vazões  $Q$  e vazão nula, respectivamente.

Além do ajuste de potência, é muito comum na literatura a utilização do ajuste polinomial de segundo grau:

$$Q = a_1 h^2 + a_2 h + a_3 \quad (4)$$

onde  $a_1, a_2$  e  $a_3$  são parâmetros de ajuste da curva.

A Figura 2 a seguir mostra um exemplo da relação cota x vazão ajustada pelas duas curvas supracitadas para o rio da Várzea no Estado do Paraná. Vale destacar que, eventualmente, encontra-se na literatura a forma gráfica invertida de curva-chave, o padrão seguido neste trabalho é baseado na escola francesa (Sefione, 2002), a qual apresenta a cota no eixo das abscissas e a vazão no eixo das ordenadas.

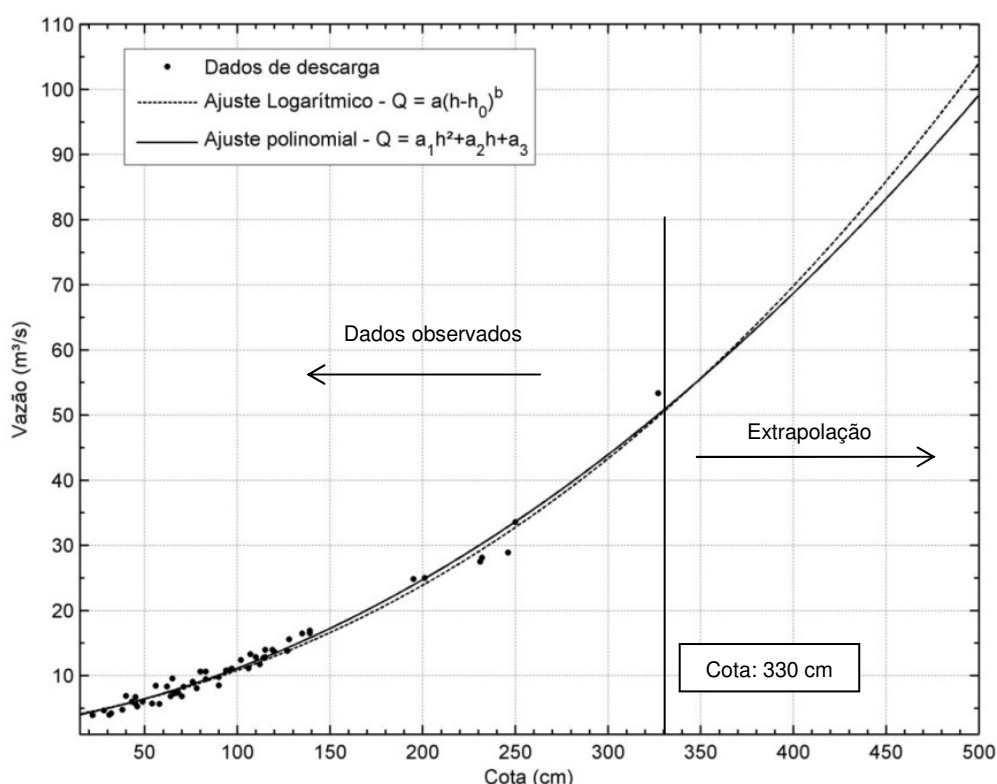


FIGURA 2: CURVA-CHAVE DA ESTAÇÃO FLUVIOMÉTRICA RIO DA VÁRZEA, ESTADO DO PARANÁ.

FONTE: O AUTOR (2014).

Outra maneira de apresentação da curva-chave é através da tabela de calibragem, na qual os valores de cota e vazão obtidos por meio das medições em campo são inseridos em duas colunas de uma tabela. Porém, a aplicação desse método tem algumas limitações: nem sempre as medições são abrangentes a ponto de não necessitar de dados intermediários, nesse caso, os valores são calculados por meio visual. Além disso, a precisão com que os instrumentos trabalham, como molinantes e ADCPs, e que os levantamentos

são feitos geram valores diferentes de vazão para uma mesma leitura de cota, sendo necessário, novamente, interpolações visuais.

Dificuldades técnicas e operacionais na medição das vazões muitas vezes impedem que os dados de descarga levantados em campo compreendam todo o período de cotas medido na estação. Dessa forma, para os valores superiores da curva-chave, onde não existe medição, é aconselhado que se use uma curva diferente do ajuste usado para os valores inferiores, a partir de informações e parâmetros possíveis de serem estimados, como a geometria da seção transversal e a rugosidade do trecho hidrométrico.

Diversos são os métodos possíveis para extrapolação de curvas-chave. No estudo de Sefione (2002) foram analisados seis métodos diferentes, sendo quatro aplicadas como sugere a bibliografia – Stevens, Manning, Manning-Fix, Velocidade-Área e Logarítmico, e dois como alternativas propostas para os métodos de Manning e Velocidade-Área, denominados no trabalho como Manning Log e VxA Log, respectivamente.

A partir dos resultados obtidos por Sefione (2002), chegou-se à conclusão de que o método de Stevens apresenta melhores resultados na extrapolação de curvas-chave. Jaccon e Cudo (1989) também citam bons ajustes da aplicação desse método para vazões máximas.

O método de Stevens está baseado na fórmula de Chezy para escoamentos uniformes em canais. Supõe-se, para uma determinada seção transversal, que a seguinte relação é válida para as vazões mais altas da curva-chave:

$$\frac{Q}{A\sqrt{R_h}} = \text{constante}. \quad (5)$$

onde  $Q$  é a vazão ( $\text{m}^3/\text{s}$ ),  $A$  é a área molhada da seção transversal ( $\text{m}^2$ ) e  $R_h$  o raio hidráulico (m). Da formulação, obtém-se que a vazão na seção é função linear de  $A\sqrt{R_h}$ . Para aumentar o grau de liberdade do ajuste, é adicionada no denominador da função mais uma constante  $d$ .

A partir dos dados de descarga e de levantamentos topográficos realizados na estação, os parâmetros que dependem da geometria da calha do rio,

como  $A$  e  $R_h$ , podem ser calculados. Adota-se, para tanto, a seguinte rotina de cálculos:

- a) Análise dos dados disponibilizados para a estação fluviométrica (cota, área molhada, seção topobatimétrica da seção transversal do rio, vazão e velocidade média do escoamento);
- b) Cálculo e determinação da seção transversal média do rio no trecho de instalação da estação fluviométrica analisada;
- c) Com os dados de área molhada levantados em campo, determina-se a profundidade da seção transversal correspondente ao valor de área levantado. A partir dele, o perímetro molhado é obtido estabelecendo um somatório da distância euclidiana entre os pontos de medição;
- d) Obtém-se as relações biunívocas lineares entre área molhada ( $A$ ) e cota ( $h$ ), raio hidráulico ( $R_h$ ) e cota ( $h$ ) e por fim, vazão ( $Q$ ) e o denominador  $A \cdot \sqrt{R_h}$  da Equação 6.

O produto resultante da aplicação do método de Stevens tem a seguinte formulação:

$$Q = A(h)\sqrt{R_h(h)}a_1 + a_2 \quad (6)$$

onde  $a_1$  e  $a_2$  são parâmetros de ajuste da relação linear. A Figura 3 mostra as relações obtidas entre os parâmetros supracitados para a estação fluviométrica Fazenda Maracanã.

Vale lembrar que para esse método ter validade uma condição torna-se necessária: o escoamento deve ser uniforme ou bem próximo disso na seção de medição. Essa consideração é razoável, uma vez que ela pode ser satisfeita se cuidados técnicos forem tomados no momento da escolha do trecho de instalação da estação (Santos *et. al.*, 2001).

Além disso, esse método só é válido para locais onde o escoamento não extravasa a calha natural menor do rio (Figura 4). Segundo os resultados obtidos por Sefione (2002), nenhum dos seis métodos propostos obteve resultados satisfatórios para extrapolação superior da curva-chave nessas regiões de várzea. Notou-se, nesses casos, que todos os métodos subestimaram a vazão da curva-chave na cota máxima.



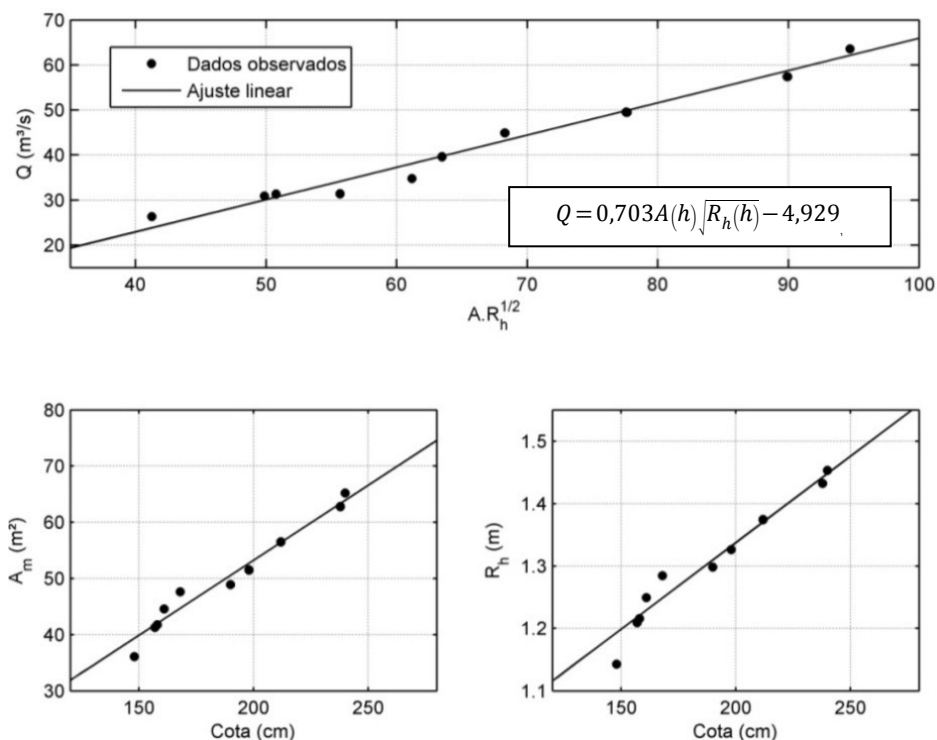


FIGURA 3: RELAÇÃO LINEAR ESTABELECIDA PELO MÉTODO DE STEVENS PARA AS VAZÕES MAIS ALTAS MEDIDAS NA ESTAÇÃO FLUVIOMÉTRICA FAZENDA MARACANÃ, ESTADO DO PARANÁ.

FONTE: O AUTOR (2014).

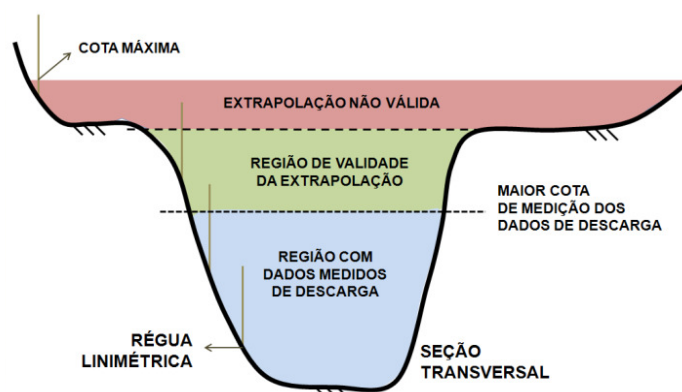


FIGURA 4: ÁREA DE VALIDADE DA EXTRAPOLAÇÃO DE CURVA-CHAVE ESTABELECIDA PELO MÉTODO DE STEVENS.

FONTE: O AUTOR (2014).

Gomes (1977) também faz uma consideração importante sobre esse método. Ele cita que a rugosidade de Manning é variável com a profundidade da calha do rio e, por isso, às vezes a aplicação do método de Stevens não se faz eficiente, uma vez que fica difícil obter a constância para toda a extensão da curva.

Antes de consistidos os dados de descarga de uma estação fluvio-métrica e gerada(s) sua(s) curva(s)-chave, Santos *et. al.* (2001) recomendam

uma análise de consistência das leituras diárias de régua a fim de detectar erros mais prováveis e corrigir possíveis incoerências.

### 3.1.2 Série histórica de leitura de régua

As leituras diárias de régua realizadas pelo operador nas estações fluviométricas têm importância fundamental na quantificação da disponibilidade hídrica da bacia de contribuição, uma vez que os dois estão diretamente relacionados.

Nem sempre as características morfológicas ou de controle se mantêm durante a vida útil de uma estação fluviométrica. Apesar de tomados os cuidados necessários na escolha do seu local de instalação, quando o monitoramento é feito por um período considerável, é possível que eventos extremos de vazão danifiquem as réguas e até mesmo modifiquem as características morfológicas da região, gerando alterações nas condições de controle da seção e conduzindo a erros sistemáticos a partir de uma determinada época.

A fim de evitar que tais erros comprometam o estudo hidrológico, é importante que seja levantada o máximo de informações possíveis a respeito da estação de medição, tais como estudos hidrológicos anteriores na mesma região, fichas descritivas e histórico, croqui de instalação da estação etc. (Tucci e Silveira, 1985).

No caso da inexistência de informações a respeito das mudanças de régua numa determinada estação, Santos *et. al.* (2001) recomendam uma análise detalhada conjunta da série histórica de dados de cota medidos durante o mais longo período possível. A Figura 5 mostra a série histórica de leituras de régua das estações ETA – Guarapuava e Ponte Pinhão, ambas localizadas na bacia hidrográfica do rio Jordão, Estado do Paraná, sendo a segunda instalada à jusante da primeira.

Apesar do órgão operador da régua (Instituto das Águas do Paraná) informar que a primeira mudança de curva-chave da estação ocorreu em dezembro de 1990, pela Figura 5 fica claro que em dezembro de 1988 houve uma mudança de régua na estação, haja vista o acréscimo repentino nos valores

das cotas que ocorreu apenas na estação ETA – Guarapuava no início desse mês.

Além da comparação entre as leituras diárias de régua, outras técnicas têm sido desenvolvidas no sentido de minimizar os erros nos estudos hidrológicos. Santos *et. al.* (2001) citam o método da curva duplo acumulativa, a qual é obtida a partir de um gráfico com as vazões (em geral médias anuais) acumuladas da estação base no eixo das ordenadas e a soma da média de várias estações vizinhas no eixo das abscissas. No caso de variação na declividade da reta de interpolação dos pontos, é provável que tenha ocorrido alguma alteração nas curvas de descarga da estação no período.

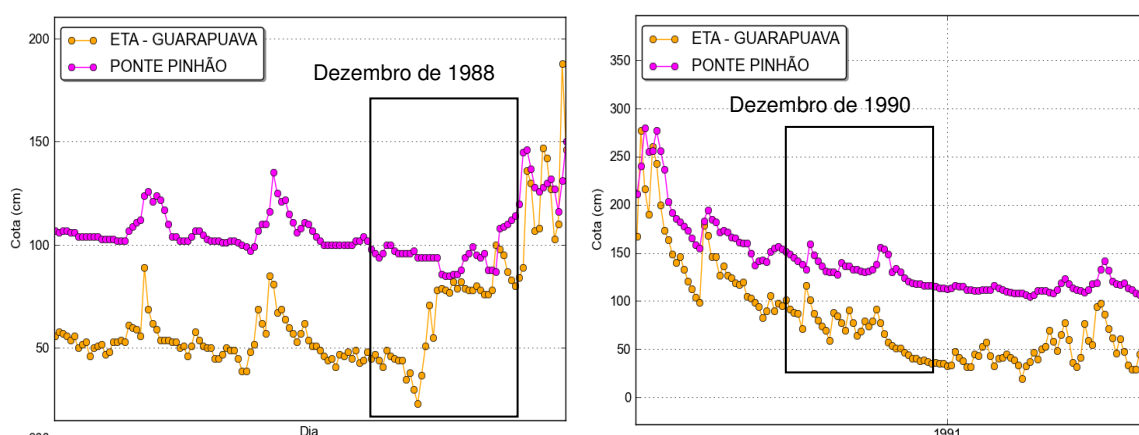


FIGURA 5: MUDANÇA DE RÉGUA OBSERVADA NA ESTAÇÃO ETA – GUARAPUAVA, ESTADO DO PARANÁ.

FONTE: O AUTOR (2014).

Outro método citado por Santos *et. al.* (2001) é o da curva de permanência adimensional. Com maior aplicabilidade para rios com mais de uma estação fluviométrica instalada, esse método se justifica porque a chance de ocorrer grandes variações nas curvas de permanência de um mesmo rio é muito pequena, mesmo aqueles com grandes variações na área de drenagem. Um ponto importante para obtenção de validade na análise é a utilização de um período comum de dados para todas as estações.

### 3.2 MODELAGEM CHUVA-VAZÃO

Os modelos precipitação-vazão representam as diversas etapas do ciclo hidrológico compreendidas entre a precipitação e a vazão. Essas etapas podem ser subdivididas em distribuição espacial da precipitação, perdas por interceptação, evaporação, acumulação em depressões do solo, fluxo da infiltração no solo, percolação e água subterrânea, escoamento superficial, sub-superficial e no rio (Tucci, 1998)

Dessa forma, geralmente modelos de simulação hidrológica podem ser aplicados para a obtenção de séries de vazões regionalizadas a partir de dados históricos de precipitação, evapotranspiração, umidade relativa do ar e até mesmo outros parâmetros que têm interferência no escoamento superficial. Para Saraiva (2010), atualmente, há uma considerável demanda em aplicar esses modelos hidrológicos para bacias hidrográficas para as quais se tem uma rede de monitoramento fluviométrico incipiente ou inexistente.

Diversas são as aplicações dos modelos de simulação hidrológica, eles podem ser utilizados tanto como mecanismo de gestão como ferramenta fundamental no planejamento de recursos hídricos. Tucci (1998) cita que os modelos chuva-vazão são utilizados não só para análise de consistência e preenchimento de falhas de medições em estações fluviométricas, como também na previsão de cenários de planejamento, dimensionamento de obras hidráulicas e até mesmo previsão de vazões em tempo real, para otimização de operação de usinas hidrelétricas.

Existem diferentes tipos de modelo, construídos com diferentes propósitos. Tucci (1998) faz uma discussão sobre aspectos relacionados à classificação de modelos.

Vendruscolo (2005) cita que normalmente os modelos são classificados de acordo com os tipos de variáveis utilizadas na modelagem, podendo ser estocásticos ou determinísticos. A maneira com as quais essas variáveis interagem pode ser empírica ou conceitual. Além disso, os dados podem ser apresentados de maneira discreta ou contínua e a existência ou não de relações espaciais caracterizam os modelos como concentrados ou distribuídos.

Geralmente, modelos conceituais tentam representar a variação das características da bacia hidrográfica no espaço a partir de balanços hídricos

localizados, determinando parcelas de evapotranspiração, infiltração, escoamento superficial e subsuperficial para cada unidade de estudo. Sendo assim, a necessidade de um levantamento mais exaustivo e detalhado é uma das maiores limitações à utilização destes modelos (Rajurkar *et. al.*, 2002). Para Tucci (1998), as incertezas de obtenção de dados de evaporação, a grande variabilidade espacial da precipitação, a não homogeneidade dos parâmetros e o fator de escala entre processos pontuais e espaciais são as principais dificuldades encontradas pelos autores e usuários desses tipos de modelo.

Por se tratar de um problema extremamente complexo e não linear, muitas vezes a modelagem chuva-vazão não pode ser realizada por simuladores que dependem do conhecimento aprofundado das características físicas da bacia hidrográfica, como é o caso dos modelos conceituais concentrados (Smap) e modelos fisicamente distribuídos (*Topmodel*). É nesse contexto que se inserem os modelos “caixa preta”, os quais funcionam de maneira mais “autônoma” e para resultados bons serem obtidos não são necessárias informações precisas de parâmetros físicos da bacia hidrográfica de aplicação, que são difíceis de serem estimados, como a taxa de infiltração do solo, a precipitação efetiva, o coeficiente de recarga do aquífero, etc.

Um exemplo destes problemas que surgem ao trabalhar com parâmetros físicos da bacia difíceis de serem mensurados foi vivenciado pelo próprio autor ao iniciar os estudos dessa dissertação, dado que lhe foi proposto trabalhar com regionalização de parâmetros do modelo chuva-vazão SMAP (*Soil Moisture Accounting Procedure*), desenvolvido por Lopes *et. al.* (1981).

Sendo assim, para esse trabalho, optou-se pela utilização do modelo empírico denominado Rede Neural Artificial (RNA). Por se tratar de um modelo potencialmente útil e amplamente aceito na modelagem de sistemas não lineares complexos com grande quantidade de dados, a utilização das RNAs na modelagem chuva-vazão tem crescido consideravelmente nos últimos anos (Maier e Dandy, 2000).

### 3.2.1 Redes Neurais Artificiais (RNAs)

A RNA é um modelo matemático fundamentado na observação e na experiência, não linear e com memória própria. Seu mecanismo de funcionamento está baseado no cérebro humano e, dependendo de sua abordagem, sua classificação pode variar entre um modelo contínuo ou discreto, concentrado ou distribuído e ainda estocástico ou determinístico (Machado, 2005).

De acordo com Valença (1999) *apud* Diniz (2008), dentre os benefícios da RNA, podem ser citados:

- Não linearidade;
- Adaptatividade e aprendizagem: dependendo da mudança de um ambiente, faz-se um novo treinamento da RNA, com aplicação de novos exemplos, e ela é capaz de se adaptar;
- Processamento paralelo: dependendo de sua arquitetura, várias são as possibilidades testadas pelas RNAs ao mesmo tempo na busca de um mínimo (ou máximo) global da função objetivo;
- Tolerância a falhas: normalmente não perde desempenho quando sujeita a condições adversas e, além disso, quando treinada adequadamente, possui uma capacidade eficiente de generalização, ao contrário de modelos físicos.

Em geral, a arquitetura de uma RNA pode ser dividida em três partes: camada de entrada, camada oculta (escondida) e camada de saída, as quais são responsáveis, respectivamente, pela captação dos dados de entrada, processamento dos dados de entrada e apresentação do resultado da rede.

Independente da arquitetura definida para uma RNA, o modelo de funcionamento citado por Warren McCulloch e Walter Pitts, em 1943, ainda é o mais utilizado, ele tem as seguintes etapas: primeiramente é dada à rede os dados de entrada  $[x_1, x_2, \dots, x_n]$ , esses são então ligados ao neurônio artificial através da sinapse que a ela é associado um peso  $[w_1, w_2, \dots, w_n]$ , em cada ligação é realizada uma operação de multiplicação. Em seguida, um somatório é estabelecido no neurônio artificial para cada entrada a ele ligado, subtrai-se então o valor do viés (*bias*)  $w_0$ . Nessa operação produz-se uma saída  $u$ , a qual é submetida a uma função de ativação  $f(u)$ , que produz uma saída  $y$ .

A Figura 6 apresenta o modelo de um neurônio artificial e a equação que governa esse modelo está apresentada na sequência.

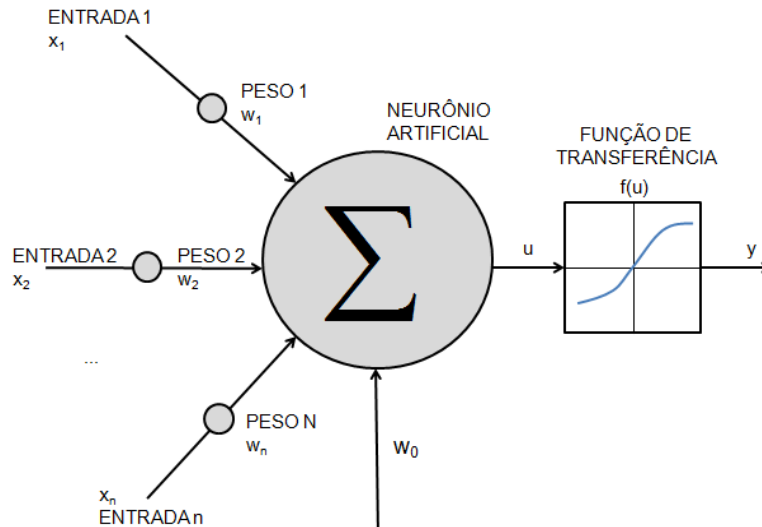


FIGURA 6: NEURÔNIO ARTIFICIAL.

FONTE: O AUTOR (2014), ADAPTADO DE MATLAB (2012).

$$y = f(u); u = \sum_{i=1}^n w_i x_i - w_0 \quad (7)$$

### 3.2.2 RNA do tipo *Multilayer Perceptron* – MLP

Uma RNA do tipo *Multilayer Perceptron* (MLP), também chamada de RNA *feedforward*, constitui-se de um conjunto de neurônios (ou nós), os quais formam as camadas de entrada, intermediária e de saída da rede. Destas camadas, a única que não é constituída por neurônios é a de entrada e, por isso, não apresenta capacidade computacional. Nesses tipos de RNAs não existem conexões diretas entre as camadas de entrada e de saída.

Na Figura 7 é apresentada uma RNA do tipo MLP formada por  $N$  entradas e  $S$  neurônios em cada camada, 2 camadas intermediárias e 1 camada de saída. Nota-se que as respostas de cada camada intermediária são entradas das próximas camadas. A fórmula que governa seu funcionamento está mostrada na equação seguinte.

$$y = f^3(CW.f^2(CW.f^1(EW.X + b^1) + b^2) + b^3) \quad (8)$$

onde  $EW$  é a matriz de pesos para os dados de entrada do vetor  $X$ , e  $CW$  representa a matriz de pesos para as camadas escondidas, na qual as linhas indicam o neurônio de destino do peso dado e a coluna indica a fonte de entrada daquele peso.

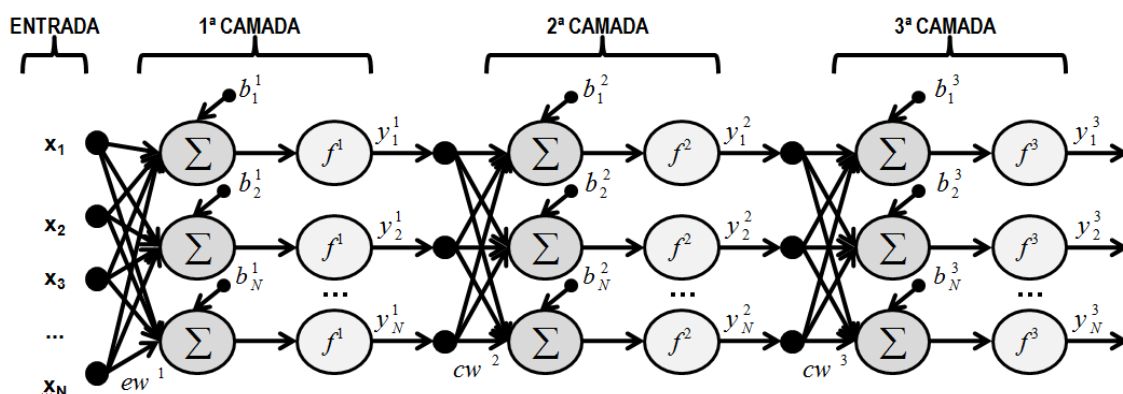


FIGURA 7: NEURÔNIO ARTIFICIAL.

FONTE: O AUTOR (2014), ADAPTADA DE MATLAB (2012).

A propagação de informação nas redes MLPs se dá sempre no sentido progressivo, camada a camada. Cada um dos neurônios da camada intermediária recebe uma combinação linear dos dados de entrada. Essa combinação gera um estímulo na função de transferência, a qual emite uma resposta. No caso das RNAs MLPs, essa resposta é tomada como entrada na camada seguinte, esta que, por sua vez, pode ser tanto mais uma camada intermediária como a própria camada de saída.

Para a solução de um determinado problema, a RNA passa por um processo de treinamento, no qual a rede criada extrai características relevantes de padrões de informações apresentadas de modo a criar uma representação própria para o problema. Essa etapa de treinamento consiste em um processo iterativo de ajuste de pesos e vieses baseado na minimização do erro na saída da RNA. Basicamente, as RNAs podem ser treinadas de maneira supervisionada e não supervisionada.

No treinamento supervisionado, o aprendizado se dá pela comparação dos dados de saída obtidos pela RNA e dos dados de saída desejados. Sempre que existir uma diferença não aceitável entre a saída obtida e a saída desejada os pesos e as conexões deverão ser reajustados até que os erros se tornem aceitáveis. Já no treinamento não supervisionado, a rede atualiza seus



pesos sem o uso de pares de entrada/saídas desejadas e sem indicações sobre adequações das saídas produzidas.

Para Souza (2012), a qualidade de uma RNA está diretamente relacionada, basicamente, a três parâmetros principais: número de padrões, tamanho e número de épocas. Machado (2005) e MATLAB (2012) citam que, além dos parâmetros acima, também devem ser considerados no momento da escolha da RNA para solução de um problema a normalização dos dados de entrada, o ajuste dos pesos, a técnica de otimização e a função de transferência.

### 3.2.2.1 Pré-processamento dos dados

Com relação ao número de padrões, os dados de entrada da rede devem refletir adequadamente a situação a ser resolvida. Nessa etapa, é importante que a amplitude dos dados cubra a variação das entradas para as quais a RNA foi treinada. MATLAB (2012) cita que as RNAs são treinadas para generalizar bem problemas, mas não são capazes de extrapolá-los.

Outro aspecto importante no pré processamento dos dados de entrada diz respeito à sua normalização. A normalização deve ser estabelecida para assegurar, no processo de treinamento, que todos os dados de entrada sejam tratados de maneira semelhante, sem favorecimento dos valores mais altos ou baixos, como aconteceria no caso de uma não normalização.

Além disso, deve haver uma compatibilidade do tipo de normalização utilizada para as variáveis com os limites das funções de ativação utilizadas na camada de saída (Minns e Hall *apud*. Maier e Dandy, 2000). Por exemplo, pelo fato das saídas da função de transferência sigmóide-logística estarem contidas no intervalo entre 0 e 1, os dados de entrada devem ser normalizados nesse mesmo intervalo, entre 0,1 e 0,9 ou 0,2 e 0,8.

Porém, quando as funções de transferência forem do tipo linear, ou seja, ilimitada, tanto para valores mais altos como para mais baixos, tal consideração não deve necessariamente ser seguida, porém recomenda-se a normalização para gerar uma uniformidade dos mesmos (Masters *apud*. Maier e Dandy, 2000).

Machado (2005) em seu trabalho faz um levantamento quanto aos tipos de normalizações geralmente utilizadas em pesquisas que envolvem RNAs. Ele cita que nos estudos de Sajikumar e Thandaveswara (1999) a normalização foi estabelecida variando os valores entre 0,05 e 0,95, seguindo a seguinte Equação:

$$x_{norm} = 0,05 + 0,90 \cdot \left( \frac{x - x_{MÍN}}{x_{MÁX} - x_{MÍN}} \right) \quad (9)$$

onde  $x$  e  $x_{norm}$  representam, respectivamente, a variável a ser normalizada e o seu valor normalizado,  $x_{MÁX}$  e  $x_{MÍN}$  são os valores máximo e mínimo da variável a ser normalizada.

Lima e Ferreira Filho (2003) utilizaram a mesma ideia, porém variando os dados de entrada entre 0,1 e 0,9, seguindo, portanto, a Equação (10).

$$x_{norm} = 0,1 + 0,8 \cdot \left( \frac{x - x_{MÍN}}{x_{MÁX} - x_{MÍN}} \right) \quad (10)$$

A mesma variação de Lima e Ferreira Filho (2003) foi utilizada por Rajurkar *et. al* (2002,2004), porém, seguindo uma equação diferente:

$$x_{norm} = 0,1 + 0,8 \cdot \left( \frac{x}{x_{MÁX}} \right) \quad (11)$$

Sperb *et. al.* (1999) em seu estudo apenas divide todos os valores de cada variável a ser normalizada pelo maior valor da variável no conjunto de entrada, conforme a Equação (12).

$$x_{norm} = \frac{x}{x_{MÁX}} \quad (12)$$

Segundo Maier e Dandy (2000), a maioria dos modelos estatísticos necessitam de um pré-tratamento dos dados de entrada de modo a torná-los normalmente distribuídos. Para o caso das RNAs, mais pesquisas devem ser realizadas para avaliar a validade dessa hipótese. Para Fortin *et. al.* (1997), devido ao fato da função erro médio quadrático ser geralmente utilizada na otimização dos pesos das RNAs, os dados de entrada devem, necessariamente, ser transformados de modo a se ajustarem à distribuição normal. Entretanto, nos testes empíricos realizados por Faraway e Chatfield (1998) essa hipótese

não foi confirmada, uma vez que os pesos atribuídos às RNAs foram os mesmos tanto para os dados transformados como para os dados brutos.

### 3.2.2.2 Arquitetura

O número de camadas, de neurônios e de sinapses é o que determina o tamanho da rede e sua arquitetura. Aumentando o número de neurônios na camada escondida aumenta-se também o poder da RNA, porém, cresce também o custo computacional e a RNA fica mais propensa a produzir o chamado *overfitting*, que é a saturação da rede com a memorização dos dados de treinamento. Dessa forma, a determinação de uma arquitetura apropriada para uma determinada RNA é uma das mais importantes e também mais difíceis etapas na construção da solução de um problema.

Maier e Dandy (2000) fazem uma abordagem bem detalhada sobre os diversos tipos de arquiteturas que vêm sendo utilizadas e suas principais vantagens e fragilidades. Os autores citam que tradicionalmente as redes *feed-forward*, onde os nós em apenas uma camada estão ligados aos nós na camada seguinte, têm sido muito utilizadas para aplicações em previsões. No entanto, as redes de repetição, as quais funcionam de maneira distinta, onde os nós de uma camada podem ser ligados aos nós tanto da próxima camada como da camada anterior, estão sendo propostos para a mesma finalidade.

Machado (2005), seguindo os resultados obtidos por Lima e Ferreira Filho (2003), trabalhou todos os modelos de RNAs com três camadas. A metodologia adotada para avaliar os resultados das diferentes arquiteturas foi a de variar o número de neurônios na camada intermediária e o número e formato das entradas na primeira camada. Ao todo foram criados 24 modelos distintos relacionados à arquitetura, as entradas variaram conforme se mostra no Capítulo 2 e os neurônios na camada intermediária variaram entre 3, 5, 8 e 10.

### 3.2.2.3 Ajuste dos pesos

A escolha da maneira como os pesos atribuídos aos dados de entrada são inicializados é fundamental para um bom ajuste da RNA, uma vez

que eles são os pontos de partida para o encontro de um mínimo global da função e, conseqüentemente, da solução de um problema.

Para HSU *et. al.* (1995), as RNAs do tipo MLP geram superfície muitas vezes irregulares e o risco de nem sempre se chegar a uma solução global é relativamente alto. Para os autores, caso o ponto inicial esteja muito próximo a um mínimo local, a solução do problema recairá próximo a esse ponto.

Geralmente, os pesos são inicializados para valores aleatórios de média nula. Maier e Dandy (2000) sugerem cautela ao definir os valores mais altos e mais baixos dos pesos  $[-w, w]$ . Caso o valor de  $w$  seja muito baixo, o treinamento pode ser paralisado, dado que a ordem de grandeza dos números fica muito pequena, mas no caso de um valor de  $w$  muito alto, a saturação prematura da RNA pode ocorrer e sua qualidade também ser reduzida (See e Shen, 1991).

Distante e Chatfield (1998) *apud.* Maier e Dandy (2000) sugerem que um certo número de diferentes conjuntos de valores iniciais aleatórios deveria ser usada para verificar se os resultados obtidos são consistentes.

Machado (2005) trabalhou com geração de números aleatórios entre -1 e 1 na inicialização dos pesos de cada RNA. Ainda assim, o autor decidiu reiniciar o treinamento de cada uma das RNAs por seis vezes, de modo a trabalhar com a mesma arquitetura pré-definida e variando apenas seus pesos iniciais a partir da geração de números aleatórios com média populacional igual a zero.

#### 3.2.2.4 Algoritmo de otimização

O algoritmo de Levenberg-Marquardt trata-se de um aperfeiçoamento do método de Gauss-Newton para solução de equações, o qual é variante do método de Newton.

Basicamente, o algoritmo de Levenberg-Marquardt utiliza informação da derivada parcial de segunda ordem do índice de desempenho utilizado para correção dos pesos, de modo a minimizar o índice e buscar o máximo (ou mínimo) global da equação (Roweis, 2003). Para fazer isso, o método tem sua

fundamentação na matriz Hessiana, a qual é definida como a matriz quadrada de derivadas parciais de segunda ordem de uma função.

Quando a função tem a forma de uma soma dos quadrados (como é o caso das RNAs *feedforward*, então a matriz de Hesse pode ser aproximada como:

$$\mathbf{H} = \mathbf{J}^T \cdot \mathbf{J} \quad (13)$$

E o gradiente calculado como

$$\mathbf{g} = \mathbf{J}^T \cdot \mathbf{e} \quad (14)$$

onde  $\mathbf{J}$  é a matriz jacobiana que contém as primeiras derivadas dos erros das redes com respeito aos seus pesos e viéses, e  $\mathbf{e}$  é um vetor de erro das redes. A matriz jacobiana pode ser calculada através de uma técnica de retro propagação, a qual é muito menos complexa do que o cálculo da matriz Hessiana.

O algoritmo de Levenberg-Marquardt utiliza essa aproximação da matriz Hessiana, resultando na seguinte expressão:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - (\mathbf{J}^T \cdot \mathbf{J} + \mu \cdot \mathbf{I})^{-1} \mathbf{J}^T \cdot \mathbf{e} \quad (15)$$

Para cada vez que o somatório do erro  $\mathbf{e}$  aumenta,  $\mu$  é multiplicado por um fator  $\beta$ , no caso do erro médio diminuir,  $\mu$  é dividido por  $\beta$ . Os autores Hagan e Menhaj (1994) recomendam os valores iniciais de  $\mu$  e  $\beta$  iguais a 0,01 e 10, respectivamente.

Reconhecido por ter uma eficiência muito alta, o algoritmo de Levenberg-Marquardt já foi bastante utilizado em estudos anteriores, como é o caso de Lima e Ferreira Filho (2003), Sperb *et. al.* (1999) e Machado (2005). MATLAB (2012) cita que o algoritmo não é recomendado para redes com milhares de pesos, pois nesses casos mais memória computacional é exigida. Além disso, o método supracitado tem bons resultados quando utilizados na solução de problemas não lineares, como é o caso do comportamento da transformação da chuva em vazão, objetivo do presente trabalho.

### 3.2.2.5 Épocas (*epochs*)

O número de épocas (*epochs*) é o número de tentativas de ajuste dos pesos e viéses da RNA para garantir um bom resultado no treinamento. Para esse parâmetro, a mesma afirmação do número de camadas é válida, o seu valor não deve ser muito elevado, nem muito pequeno, pois podem ocorrer, respectivamente, *overfitting* ou incapacidade de memorização de padrões fornecidos.

### 3.2.2.6 Função de transferência

A função de transferência, denotada por  $f()$  e também chamada de função de ativação, tem como característica restringir a amplitude da saída de um neurônio, uma vez que limita o intervalo de amplitude do sinal de saída a um valor finito. Geralmente esse intervalo está entre 0 e 1, ou entre -1 e 1.

Dentre os diversos tipos de função de transferência, podem ser citadas as: função de transferência linear, função de transferência sigmoide logística, sigmoide tangente hiperbólica e degrau, mostradas na Figura 8.

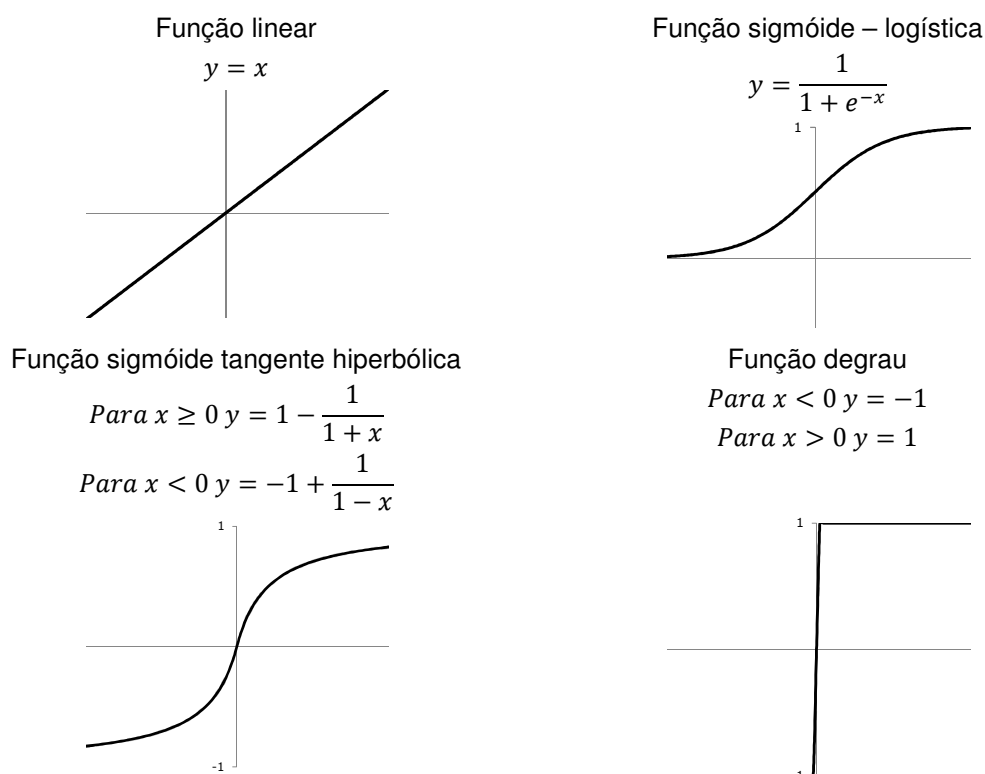


FIGURA 8: FUNÇÕES DE TRANSFERÊNCIA  
 FONTE: ADAPTADO DE MACHADO (2005)

Em problemas de recursos hídricos, as funções de transferência mais utilizadas são a sigmóide – logística na camada intermediária e linear na camada de saída (Rajurkar *et. al.*, 2002). Tal afirmação é coerente, uma vez que as respostas da função sigmóide - logísticas são todas positivas, o que fisicamente faz sentido em se tratando de processos hidrológicos.

### 3.3 INTERPOLAÇÃO ESPACIAL

Interpolação espacial é o processo de estimar o valor de propriedades em locais não amostrados localizados dentro de áreas cobertas por observações existentes. Para Stark (1979), interpolar nada mais é do que o processo matemático de achar valores intermediários entre os valores discretos de uma função.

A interpolação espacial pode ser utilizada na transferência de dados e no planejamento de redes de estações para coleta de dados. Nos estudos de Kaviski (1992) é feita uma abordagem geral sobre as diversas técnicas utilizadas no mapeamento regional baseado em interpolações espaciais.

Dentre as mais frequentes, podem ser citados os mapas de isolinhas e poligonais, que no ramo da hidrologia são utilizados tanto no cálculo de média de precipitação acumulada em bacias hidrográficas como na estimativa de valores de temperatura em locais sem postos de medição.

Totais de precipitação média obtidos a partir de um conjunto de postos de medição localizados no interior (ou próximos ao divisor de águas) de bacias hidrográficas são parâmetros frequentes analisados em estudos hidrológicos. Para o cálculo, a maioria dos procedimentos que são utilizados pode ser expresso como uma combinação linear de todas as observações feitas nas estações.

Para Maidment (1992), se  $n$  postos de medição, com valores de  $P_1$ ,  $P_2, \dots, P_n$ , estão disponíveis para estimativa de precipitação em uma determinada região, então uma estimativa para a precipitação média nessa área pode ser dada pela média ponderada:

$$\bar{P} = \sum_{i=1}^n a_i P_i \quad (16)$$

onde os pesos dados às estações  $a_1, a_2, \dots, a_n$  são constantes positivas e seu somatório deve ser igual à unidade.

O método dos polígonos de Thiessen consiste basicamente na subdivisão da área da bacia em sub-regiões formadas pelo encontro das mediatrizes das retas que interligam as estações pluviométricas inseridas na região. Esse método considera que dentro do polígono de influência de uma determinada estação o total de precipitação medida por ela é constante.

Para o cálculo da precipitação média na bacia, o peso dado a cada dado de medição é ponderado pela área de abrangência do polígono formado pela união das retas.

Muitas vezes é conveniente para o cálculo da precipitação média na bacia a subdivisão da mesma em compartimentos retangulares infinitesimais, os quais terão seus valores estimados através de técnicas de interpolação espacial. O método do inverso do quadrado da distância é comumente utilizado para esse propósito. A estimativa do valor para um determinado retângulo é dada por:

$$\bar{P}_j = a \sum_{i=1}^n d_{ij}^{-2} P_i \quad (17)$$

onde  $d_{ij}$  é a distância desde a estação fluviométrica  $i$  até o centro do retângulo  $j$  e  $a$  é o inverso da soma do método do inverso do quadrado da distância, dado por:

$$a = \left( \sum_{i=1}^n d_{ij}^{-2} \right)^{-1} \quad (18)$$

Considerando que a bacia seja subdividida em  $n$  subregiões (retângulos), a sua média de precipitação será obtida pela média aritmética dos  $n$  retângulos formados em seu interior.

Na Figura 9 está mostrada a divisão da bacia hidrográfica do rio Pirapó pelas áreas de influências das estações pluviométricas localizadas no seu



entorno, utilizando o método de Thiessen e do inverso do quadrado da distância para o cálculo das isolinhas.

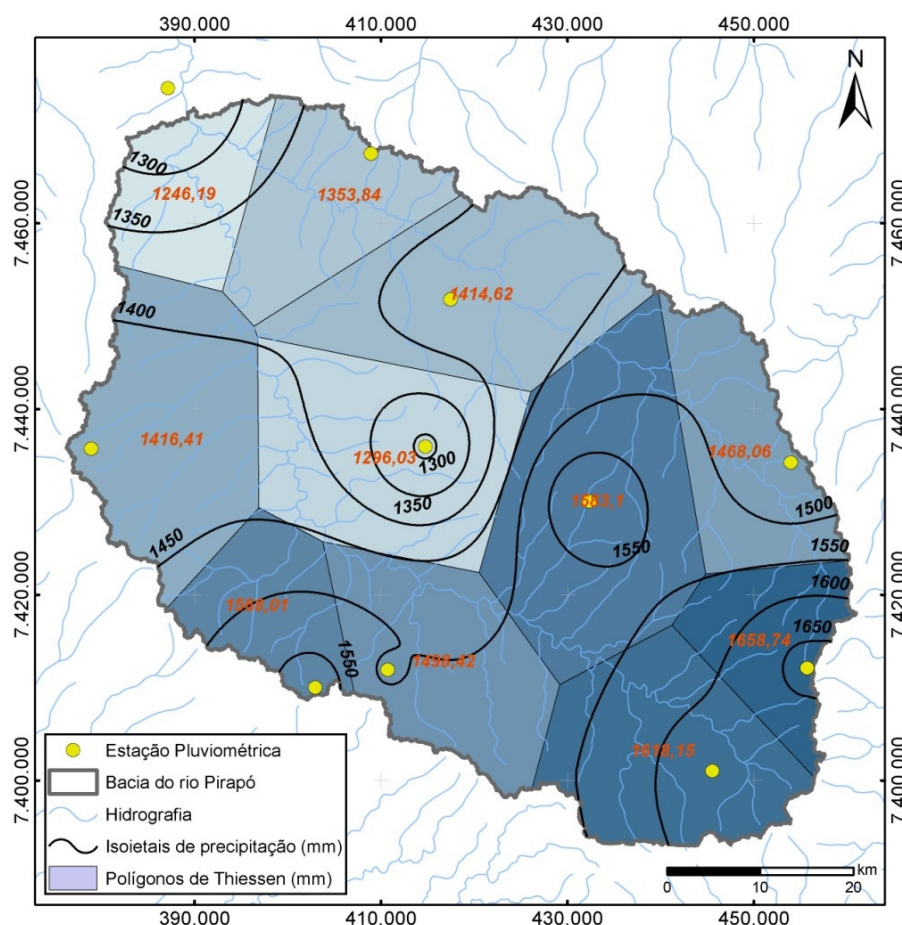


FIGURA 9: MÉTODOS DE THIESSEN E DAS ISOIETAS (INVERSO DO QUADRADO DA DISTÂNCIA) APLICADO À BACIA HIDROGRÁFICA DO RIO PIRAPÓ, ESTADO DO PARANÁ.

FONTE: O AUTOR (2014).

Singh e Chowdhury (1986) *apud*. Maidment, (1992) estudaram o desempenho de vários métodos para a estimativa de precipitação média em bacias e mostraram que todos os métodos apresentam bons resultados quando comparados por um longo período de tempo. Porém, conforme o período de análise é reduzido, as diferenças entre os resultados dos diferentes métodos aumentavam.

### 3.4 ANÁLISE ESTATÍSTICA MULTIVARIADA

Ferreira (2008) faz uma breve introdução à estatística multivariada em seu livro, para ele

“A multivariada é uma fascinante área da estatística que estuda os fenômenos observando e analisando suas diversas respostas simultaneamente. (...) A estatística multivariada tem como ingrediente básico a resposta de várias variáveis nas unidades amostrais ou experimentais. A análise, a descrição e a inferência são realizadas com base nas respostas simultâneas, valendo-se da estrutura de correlação entre as variáveis. Ademais, a estatística multivariada possui técnicas inferenciais nas mais variadas situações contempladas pelo seu contraposto univariado.”

A estatística multivariada permite o estudo de fenômenos complexos, pois realiza o tratamento de diversas variáveis simultaneamente, mesmo quando não se conhece o modelo teórico das relações entre as variáveis.

Em hidrologia, por conta das diversas variáveis que têm influência no balanço hídrico de uma bacia hidrográfica, a utilização de métodos estatísticos multivariados tem ganhado destaque nas últimas décadas. Como já citado no Capítulo 2, Kaviski (1992) e CEHPAR (2006) utilizaram muito tais técnicas para regionalização de vazões no Estado do Paraná na década de 1990.

#### 3.4.1 Análise de componentes principais

A análise de componentes principais (ACP) consiste em uma técnica estatística multivariada que tem por objetivo principal a redução da dimensão dos dados descritivos de um problema mantendo o grau de significância de toda a amostra.

Matematicamente, a ACP pode ser entendida como uma transformação linear ortogonal que substitui os dados do conjunto de variáveis originais por um novo conjunto de variáveis que, mantém, ao máximo, a variabilidade do conjunto. Em outras palavras, é um método utilizado para transformar variáveis correlacionadas em variáveis não correlacionadas, servindo ainda na obtenção de combinações lineares das variáveis originais com variabilidade relativamente grande (ou pequena, dependendo do caso) (Moita, 2004 *apud*. Smaka, 2010).

Dadas múltiplas observações de um vetor  $x$  ( $K \times 1$ ), a ACP encontra ( $M \times 1$ ) vetores  $u$  cujos elementos são combinações lineares dos elementos  $x$ 's, que contém o máximo de informação dos dados originais de  $x$ . Tal análise

é mais efetiva quando o número de variáveis a ser reduzida à amostra é consideravelmente menor do que a quantidade de dados originais (Wilks, 2006).

As novas variáveis ou componentes principais  $\mathbf{u}_m$  de  $\mathbf{u}$  são definidos pelos autovetores da matriz covariância (ou correlação) de  $\mathbf{x}$ . Particularmente, o m-ésimo componente principal  $\mathbf{u}_m$  é obtido como a projeção dos dados do vetor  $\mathbf{x}'$  no m-ésimo autovetor em:

$$\mathbf{u}_m = \mathbf{e}_m^T \mathbf{x}' = \sum_{k=1}^K e_{km} \mathbf{x}'_k \text{ onde } m = 1, \dots, M. \quad (19)$$

Nota-se que cada um dos  $M$  autovetores corresponde à uma base para cada uma das  $K$  variáveis,  $\mathbf{x}_k$ . Isso é, os autovetores definem um novo sistema de coordenadas para visualização dos dados, o qual é orientado de modo que cada um deles seja direcionado para o máximo de variabilidade dos dados, mantendo sempre a ortogonalidade.

A principal propriedade dos componentes principais é que eles são não correlacionados. Isso é, a matriz de correlação das novas variáveis  $\mathbf{u}_m$  é igual à matriz identidade  $\mathbf{I}$ . Além disso, a variância individual explicada por cada componente principal equivale ao valor do autovalor correspondente àquele componente.

$$VAR(\mathbf{U}_m) = \mathbf{e}_m' \Sigma \mathbf{e}_m = \lambda_m \quad (20)$$

Onde  $\Sigma$  é a matriz de covariância associada com o vetor aleatório  $\mathbf{u}_m$ .

Conclui-se, portanto, que a proporção da variância total devida à m-ésima componente principal é:

$$\%VAR.EXP(\mathbf{u}_m) = \frac{\lambda_m}{\lambda_1 + \lambda_2 + \dots + \lambda_k} \times 100 \quad (21)$$

Não existe definida uma regra final que determine o número de componentes que devem ser selecionados para a explicação de um fenômeno. Dentre os critérios geralmente utilizados merece destaque aquele que estabelece a permanência das  $n$  primeiras componentes ( $n < p$ ), de acordo com a variação cumulativa que elas explicam. Uma variação cumulativa com um bom grau de explicação é superior a 75%, porém, fica a cargo do usuário julgar a necessidade de variação a ser explicada.

Existem duas interpretações diferentes que podem ser feitas ao estudar os resultados da ACP. Uma delas é analisar as variáveis obtidas e trabalhar com os componentes principais no novo sistema de coordenadas criado, esse método faz com que a nova amostra não tenha mais significado físico, mas sim matemático.

O outro método consiste em avaliar quais são os indivíduos  $x_m$  que, de fato, têm influência sobre o problema analisado. Para isso, basta ordenar os autovalores e autovetores correspondentes em ordem decrescente de significância. A partir de então, busca-se nos autovetores a linha  $n$  que detém o maior valor absoluto dentre todas as linhas do mesmo componente principal, essa linha  $n$  está relacionada ao  $n$ -ésimo indivíduo analisado e, portanto, esse indivíduo explica o fenômeno.

### 3.4.2 Análise de agrupamentos

A análise de agrupamentos é uma ferramenta importante na classificação de objetos, itens ou indivíduos de acordo com suas similaridades a fim de classificar grupos homogêneos. Os indivíduos (objetos ou itens) similares são alocados em um mesmo grupo e, dessa forma, aqueles que pertencem a diferentes grupos são considerados dissimilares.

Os métodos de agrupamentos são divididos, de uma maneira geral, em hierárquicos e não hierárquicos. O que diferencia um do outro, praticamente, é a maneira de abordagem do problema. No caso dos métodos hierárquicos, parte-se de  $n$  grupos, cada um contendo um dos objetos e após os devidos agrupamentos todos os objetos compõem um único grupo. Tal método é denominado hierárquico aglomerativo. Ainda nos métodos hierárquicos, existe a subdivisão do hierárquico divisivo, no qual a análise inicia-se pela formação de um único grupo, o qual vai se ramificando até formação dos  $n$  grupos com um único objeto cada.

No caso dos agrupamentos não hierárquicos, a primeira etapa é a definição do número  $k$  de grupos, para então alocar os  $n$  objetos aos  $k$  grupos de maneira otimizada. Para isso, torna-se necessário primeiramente uma alo-

cação arbitrária para ser então inicializado processo iterativo buscando a alocação ótima.

Dentre os métodos de agrupamentos não hierárquicos, o mais usado é o método das k-médias (Ferreira, 2008). O mesmo autor sugere a implementação do seguinte algoritmo para utilização do método das k-médias:

- a) Alocar arbitrariamente os  $n$  objetos aos  $k$  grupos e calcular os seus centróides;
- b) Alocar cada um dos  $n$  objetos aos grupos que apresentam a menor distância com o respectivo objeto. A distância entre o objeto e o grupo é obtida, em geral, pela distância entre o vetor de observação do objeto e o centróide do grupo;
- c) Recalcular os centróides de cada grupo;
- d) Realocar o primeiro objeto de seu grupo para um outro grupo em que a distância for mínima e, obviamente, menor do que a distância desse objeto para o seu próprio grupo de origem;
- e) Repetir os passos c) e d) até que não ocorram mais mudanças de objetos de um grupo para outro. Nesse processo é realizada apenas uma transferência por iteração.

Em ambiente MATLAB, a função *kmeans* usa um algoritmo iterativo que minimiza a soma das distâncias de cada objeto com o centróide de cada grupo, sobre todos os grupos. O resultado é a formação de grupos mais compactos e mais bem separados possível.

Dos problemas que merecem destaque ao estudar métodos não hierárquicos está a definição do número de grupos a serem formados pelos  $n$  indivíduos. O método das silhuetas (MATLAB, 2012) ajuda na solução desse problema.

Após formados todos os grupos, o valor silhueta para o  $i$ -ésimo indivíduo  $S_i$  é definido como:

$$S_i = \frac{b_i - a_i}{\max(a_i, b_i)} \quad (22)$$

onde  $a_i$  é a distância média do  $i$ -ésimo indivíduo aos outros indivíduos do mesmo grupo, e  $b_i$  é a distância mínima entre o  $i$ -ésimo indivíduo para os indivíduos dos diferentes grupos. Quanto mais próximos da unidade são os valores

de  $S_i$  obtidos para os indivíduos, mais próximos eles estão com aqueles que foram o seus próprios grupos e mais distantes eles estão dos outros grupos, indicando grande proximidade interna e considerável dissimilaridade externa. A Figura 10 apresenta um gráfico de silhueta para um agrupamento em 4 grupos realizados a partir de dados disponíveis no MATLAB.

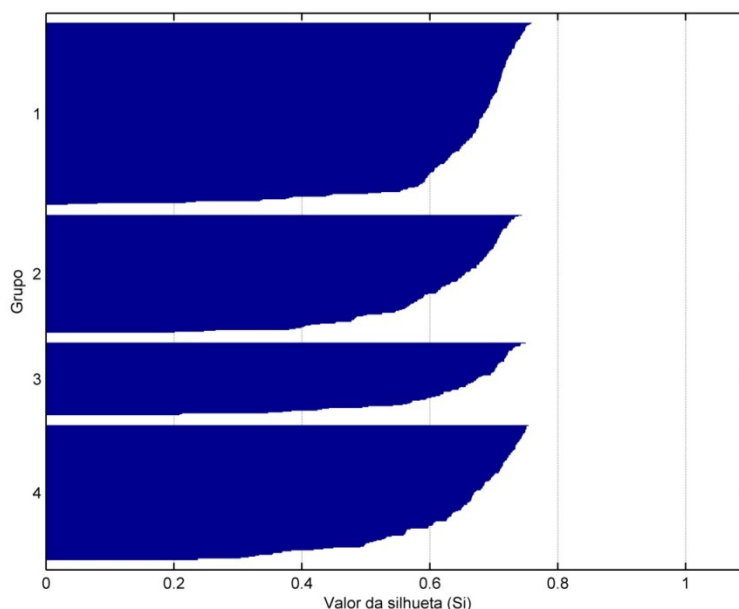


FIGURA 10: VALORES DA SILHUETA PARA UM AGRUPAMENTO NÃO HIERÁRQUICO DE DADOS.

FONTE: O AUTOR (2014) ADAPTADO DE MATLAB (2012).

### 3.4.3 Análise discriminante

A análise discriminante trata-se de uma técnica estatística multivariada utilizada para a alocação de uma amostra ou população em grupos nos quais, *a priori*, tem-se conhecimento de suas características gerais. Seu uso é amplo e o método se aplica a diversas áreas do conhecimento, como na classificação de insetos na ecologia, na análise de créditos em finanças, no diagnóstico de áreas de maior potencial para plantação de determinadas sementes na agricultura, entre outros (Mingoti, 2005).

Matematicamente esta técnica consiste em encontrar uma combinação ou transformação de características que minimiza a relação da diferença entre as médias multivariadas do grupo de características da bacia e a variância multivariada das características da bacia entre os grupos.

Seu objetivo, portanto, é o de construir uma regra de classificação que minimize tanto o número de classificações incorretas como também o cus-

to desta classificação incorreta. Para isso, a regra discriminante de máxima verossimilhança para alocação das observações  $x$  para um dos grupos  $G^{(k)}$  ( $1 \leq k \leq K$ ) é alocar  $x$  para o grupo  $G^{(j)}$  para o qual resulte na maior verossimilhança para  $x$ , ou seja, quando a equação a seguir seja atendida:

$$L_j(x) = \max(L_k(x)) \quad (23)$$

onde  $L_k(x)$  é a função de verossimilhança do  $k$ 'ésimo grupo.

Supondo-se que em cada grupo  $G^{(k)}$  as amostras sejam distribuídas normalmente com parâmetros  $\pi_k$  e  $\Sigma$ ,  $k = 1, \dots, K$ , com  $\Sigma > 0$ , então a regra discriminante de máxima verossimilhança aloca  $x$  para a  $G^{(j)}$  quando  $j$  é o valor de  $k$ , tal que é mínimo o quadrado da distância de Mahalanobis, dada por:

$$D_k(x) = (x - \mu_k)^T \Sigma^{-1} (x - \mu_k) \quad (24)$$

Se os grupos tem probabilidade *a priori*  $\pi_1, \dots, \pi_k$ , então a regra discriminante de Bayes com respeito a  $\pi$  aloca a observação  $x$  para o grupo em que  $n_k L_k(x)$  é minimizada. Neste caso, aloca-se  $x$  para  $G^{(j)}$ , quando  $j$  é o valor de  $k$ , tal que é mínima a seguinte expressão:

$$D_k(x) = \frac{(x - \mu_k)^T \Sigma^{-1} (x - \mu_k)}{2} - \ln(\pi_k) \quad (25)$$

Tasker (1982) *apud* Kaviski (1992) assumiu que as probabilidades *a priori* são proporcionais ao número de estações  $n_k$  com dados que define cada grupo, desta forma, a expressão anterior se reduz a:

$$D_k(x) = \frac{(x - \mu_k)^T \Sigma^{-1} (x - \mu_k)}{2} - \ln(n_k) \quad (26)$$

Para Kaviski (1992), a aplicação de técnicas de análise discriminante em estudos de regionalização é imediata quando os indivíduos a serem classificados são bacias hidrográficas sem dados hidrológicos, representadas por características climáticas e/ou fisiográficas, e os grupos homogêneos forem definidos em relação a estas mesmas características com base nos dados das estações hidrológicas com dados hidrológicos observados.

#### 4 APLICAÇÃO DO MÉTODO PROPOSTO AO ESTUDO DE CASO

O estudo de caso foi desenvolvido para a região centro-oeste do Estado do Paraná, a qual, depois da Região Metropolitana de Curitiba, apresenta uma maior densidade de estações de monitoramento fluviométrico e pluviométrico e ao mesmo tempo uma boa distribuição geográfica, com variações consideráveis de altitude, declividade, textura de solo e geologia, critérios importantes no momento da realização da regionalização.

Além destes, estão apresentados na sequência os principais critérios adotados para a inclusão das bacias hidrográficas no presente estudo:

- Área de drenagem da estação fluviométrica com valor inferior a 5.000 km<sup>2</sup>, uma vez que todas as áreas de drenagem das estações utilizadas no estudo estão compreendidas neste intervalo;
- Disponibilidade de dados confiáveis e representativos, de vazão e precipitação, para o período de janeiro de 1998 a dezembro de 2011, totalizando 16 anos de medição, com períodos de falhas não superiores a 5 meses sequenciais;
- Bacias hidrográficas com pequena expansão urbana em seu interior dentro do período de 1998 a 2011, haja vista a influência da impermeabilização do solo no escoamento superficial e;
- Curso d'água não afetado por desvios ou regularização de vazão à montante da régua linimétrica.

Para cada estação fluviométrica analisada, as seguintes etapas foram seguidas:

- a) Traçado de sua área de drenagem e cálculo de seu centroide;
- b) Tratamento, análise de consistência e geração de sua série histórica de vazão média mensal;
- c) Tratamento, análise de consistência e geração da série de precipitação mensal da bacia hidrográfica. Como produtos foram obtidos tanto a série histórica mensal como a precipitação média da bacia gerada pelo método de Thiessen (citado no item 3.3);



- d) Tratamento, análise de consistência e geração da série histórica das variáveis meteorológicas (temperatura, umidade relativa e radiação solar) a partir da interpolação, pelo método do inverso do quadrado da distância, aplicada no centroide da bacia hidrográfica;
- e) Treinamento e validação de uma RNA para cada estação fluviométrica;

Após isso, serão selecionadas duas estações teste, as quais servirão para avaliar a validade da regionalização das RNAs. A escolha da RNA a ser utilizada em cada estação será definida a partir de técnicas de análise multivariada (análise de agrupamentos e análise discriminante).

No desenvolvimento do trabalho foram utilizados os softwares ArcGis 9.3 e Global Mapper v.14 para o mapeamento e SIG. O MATLAB R2009a, por meio de sua biblioteca de RNAs, foi utilizado para a criação e validação das RNAs.

#### 4.1 DISPONIBILIDADE DE DADOS

Os dados hidrometeorológicos utilizados para o desenvolvimento do estudo foram provenientes de órgãos federais e estaduais. A Agência Nacional de Águas - ANA, através do site “HIDROWEB”, e o Instituto das Águas do Paraná disponibilizaram os seguintes dados referentes à fluviometria e pluviometria:

- Histórico de medições de cotas das estações fluviométricas;
- Resumo de descarga líquida;
- Levantamentos dos perfis transversais das seções dos rios e;
- Série histórica de precipitação das estações.

A Tabela 2 mostra todas as estações fluviométricas utilizadas no presente estudo. Na última coluna desta tabela estão apresentadas as estações pluviométricas inseridas, ou bem próximas, da área de drenagem de cada uma das estações fluviométricas, mostrada na terceira coluna. O detalhamento das estações pluviométricas está apresentado na Tabela 3.

TABELA 2: ESTAÇÕES FLUVIOMÉTRICAS E PLUVIOMÉTRICAS SELECIONADAS PARA O ESTUDO.

Código ANA	Estação Fluviométrica	Área de drenagem (km <sup>2</sup> )	Curso d'água	Estações pluviométricas
64775000	Balsa do Cantú	2520,1	Rio Cantú	Altamira do Paraná, Balsa do Cantú, Barra Grande, Laranjal, Palmital e Roncador
65809000	ETA - Guarapuava	312,8	Rio das Pedras	Guarapuava, Campo de Dentro, Invernadinha e Palmeirinhas
65415000	Fazenda Maracanã	323,6	Rio Palmital	Fazenda Maracanã, Pátio Velho, Salto do Vau e Santana
64764000	Guampará	1688,7	Rio Piquiri	Campina do Simão, Marquinhos, Palmital e Santa Maria do Oeste
65764000	Madeira Gavazzoni	935,7	Rio da Areia	Campina Bonita, Colônia Alegrave, Fazenda Zaniolo, Inácio Martins, Leonópolis e Madeira Gavazzoni
64785000	Ponte do Goio Bang	1341,8	Rio Tricolor	Campina da Lagoa, Guarani, Mamborê, Pensamento e Ponte do Goio-Bang
64773000	Ponte Leôncio Primo	696,6	Rio Cantú	Barra Grande, Roncador, Palmital e Pitanga
64767000	Porto Carriel	3569,3	Rio Piquiri	Campina do Simão, Campo Verde, Laranjal, Marquinhos, Palmital, Porto Carriel e Santa Maria do Oeste
64771500	Porto Guarani	4165,0	Rio Piquiri	Campina do Simão, Campo Verde, Laranjal, Marquinhos, Palmital, Porto Carriel e Santa Maria do Oeste
65970001	Porto Santo Antônio	1084,1	Rio Guarani	Goioxim, Guaraniaçu, Linha Mirim, Porto Santo Antônio, Torra Alta e Vila Guarani
64620000	Rio dos Patos	1097,7	Rio dos Patos	Apiaba, Itapara, Relógio, Rio dos Patos, Tijuco Preto e Inácio Martins
64790000	Salto Sapucaí	693,3	Rio Sapucaí	Corbélia, Ponte Tourinho – Braganey, Salto Sapucaí e São João do Oeste
65979000	São Sebastião	1311,6	Rio Andrada	Boi Preto, Cascavel – Ocepar, Flor da Serra, Linha Santa Catarina, Nelson Taborda, Rio do Salto, Santa Lúcia e São João do Oeste
64625000	Tereza Cristina	3567,9	Rio Ivaí	Apiaba, Bairro da Imbuia, Campo de Dentro, Inácio Martins, Itapara, Laranjeiras, Relógio, Rio dos Patos, Saltinho, Tereza Cristina e Tijuco Preto
65971050	Tormenta Fazenda Trento	319,6	Rio Tormen-ta	Catanduvas, Flor da Serra, Ibema, Ibira-cema, Nelson Taborda e Rio do Salto

FONTE: ANA (2012), AGUASPARANA (2012), O AUTOR (2014).

O órgão responsável pelo fornecimento dos dados meteorológicos foi o Instituto Tecnológico Simepar - SIMEPAR, que disponibilizou a série histórica de medição das seguintes variáveis levantadas em todos os postos de monitoramento inseridos no Estado do Paraná: temperatura, temperatura máxima, temperatura mínima, umidade relativa e radiação solar. A Tabela 4, a seguir, mostra todas as estações meteorológicas utilizadas no presente estudo.

TABELA 3: ESTAÇÕES PLUVIOMÉTRICAS UTILIZADAS NO ESTUDO.

Código ANA	Estação	Município	Latitude	Longitude	Altitude (m)
2651016	Salto do Vau	União da Vitória	-26:3:0	-51:12:0	800
2551017	Santana	Cruz Machado	-25:56:38	-51:15:44	880
2551000	Guarapuava	Guarapuava	-25:27:0	-51:27:0	950
2551034	Invernadinha	Guarapuava	-25:18:0	-51:26:0	1050
2551033	Palmeirinha	Guarapuava	-25:15:0	-51:33:0	1050
2451013	Pitanga	Pitanga	-24:45:0	-51:46:0	860
2553026	Flor da Serra	Boa Vista da Aparecida	-25:23:40	-53:23:5	478
2553031	Ibiracema	Cascavel	-25:17:0	-53:19:0	640
2553040	Catanduvas	Catanduvas	-25:12:0	-53:8:0	650
2553016	Nelson Taborda	Cascavel	-25:10:0	-53:22:0	828
2553035	Rio do Salto	Cascavel	-25:7:59	-53:19:0	606
2553019	Ibema	Ibema	-25:6:0	-53:4:0	750
2551023	Fazenda Zaniolo	Cruz Machado	-25:49:0	-51:6:0	1065
2551022	Madeira Gavazzoni	Inácio Martins	-25:48:27	-51:17:19	872
2551001	Leonópolis	Inácio Martins	-25:41:39	-51:12:9	960
2551011	Colônia Dalegrave	Inácio Martins	-25:37:47	-51:5:16	1150
2551026	Campina Bonita	Inácio Martins	-25:34:0	-51:15:0	1108
2551014	Inácio Martins	Inácio Martins	-25:34:0	-51:4:0	1268
2553056	Porto Santo Antônio - 396	Três Barras do Paraná	-25:23:43	-53:6:16	430
2552046	Linha Mirim	Quedas do Iguaçu	-25:23:34	-52:57:52	666
2553041	Torra Alta	Catanduvas	-25:13:0	-53:3:0	600
2552047	Vila Guarani	Nova Laranjeiras	-25:12:41	-52:39:6	917
2552026	Goioxim	Goioxim	-25:11:59	-51:59:27	858
2552006	Guaraniaçu	Guaraniaçu	-25:5:0	-52:53:0	920
2651056	Palmital do Meio	União da Vitória	-26:1:50	-51:8:31	840
2551018	Pátio Velho (Pinaré)	Cruz Machado	-25:54:11	-51:6:16	1114
2551014	Inácio Martins	Inácio Martins	-25:34:0	-51:4:0	1268
2550053	Itapara	Irati	-25:28:0	-50:59:0	900
2551038	Tijuco Preto	Prudentópolis	-25:22:47	-51:4:58	792
2550043	Apiaba	Imbituva	-25:20:0	-50:46:0	900
2551039	Relógio	Prudentópolis	-25:17:0	-51:6:0	850
2551035	Campo de Dentro	Guarapuava	-25:16:0	-51:15:0	1202
2550000	Rio dos Patos	Prudentópolis	-25:12:0	-50:56:0	690
2450049	Saltinho	Ivaí	-24:57:27	-50:53:30	743
2451035	Bairro da Imbuia	Reserva	-24:51:33	-51:0:32	934
2451003	Tereza Cristina	Cândido de Abreu	-24:50:0	-51:9:0	550
2451039	Laranjeiras	Laranjeiras do Sul	-25:25:0	-52:25:0	880
2452001	Ponte do Goio - Bang	Campina da Lagoa	-24:37:0	-52:56:0	320
2452011	Campina da Lagoa	Campina da Lagoa	-24:35:59	-52:48:15	618
2452035	Guarani	Mamborê	-24:25:59	-52:33:0	650
2452040	Pensamento	Mamborê	-24:19:0	-52:39:0	638
2452014	Mamborê	Mamborê	-24:16:59	-52:31:0	702
2453056	São João do Oeste	Cascavel	-24:57:46	-53:14:38	697
2453014	Ponte Tourinho - Braganey	Campo Bonito	-24:52:59	-53:4:0	700
2453012	Corbelia	Corbélia	-24:48:0	-53:18:0	682

continua

continuação

Código ANA	Estação	Município	Latitude	Longitude	Altitude (m)
2453013	Salto Sapucaí	Iguatu	-24:38:0	-53:6:0	730
2553028	Santa Lúcia	Santa Lúcia	-25:24:0	-53:34:0	380
2553026	Flor da Serra	Boa Vista da Aparecida	-25:23:40	-53:23:5	478
2553029	Linha Santa Catarina	Santa Lúcia	-25:21:0	-53:33:0	344
2553016	Nelson Taborda	Cascavel	-25:10:0	-53:22:0	828
2553033	Boi Preto	Santa Tereza do Oeste	-25:9:0	-53:37:0	668
2553035	Rio do Salto	Cascavel	-25:7:59	-53:19:0	606
2453056	São João do Oeste	Cascavel	-24:57:46	-53:14:38	697
2453023	Cascavel - Ocepar	Cascavel	-24:56:0	-53:26:0	760
2451021	Barra Grande	Santa Maria do Oeste	-24:46:45	-51:57:16	929
2552019	Campo Verde (Faxinal dos Índios)	Laranjeiras do Sul	-25:13:21	-52:25:52	785
2552008	Marquinhos	Marquinho	-25:6:44	-52:15:30	872
2551009	Campina do Simão	Guarapuava	-25:6:33	-51:48:23	1056
2452051	Porto Carriel	Diamante do Sul	-24:57:0	-52:37:0	380
2451010	Santa Maria do Oeste	Santa Maria do Oeste	-24:56:0	-51:53:0	900
2452019	Laranjal	Laranjal	-24:53:9	-52:28:26	741
2452016	Palmital	Palmital	-24:53:4	-52:12:10	890
2452019	Laranjal	Laranjal	-24:53:9	-52:28:26	741
2452016	Palmital	Palmital	-24:53:4	-52:12:10	890
2452012	Altamira do Paraná	Altamira do Paraná	-24:48:0	-52:42:0	650
2451021	Barra Grande	Santa Maria do Oeste	-24:46:45	-51:57:16	929
2452000	Balsa do Cantu	Altamira do Paraná	-24:44:56	-52:42:7	360
2452015	Roncador	Roncador	-24:36:0	-52:16:0	730

FONTE: O AUTOR ADAPTADO DE ANA (2012).

Todas as estações hidrometeorológicas analisadas, assim como a área de abrangência do presente estudo, estão mostradas na Figura 11.

Pelo fato do início da operação das estações meteorológicas ter-se dado praticamente na metade ano de 1997, a série histórica tomada como padrão no presente estudo envolveu os anos de 1998 a 2011, totalizando 16 anos completos de amostra, avaliados em escala mensal.

## 4.2 TRATAMENTO DOS DADOS

Nem todas as estações de monitoramento apresentaram a série completa de medições. Dessa maneira, técnicas diferentes de consistência simples foram utilizadas de modo a completar tais períodos faltantes e compor uma série histórica mensal.

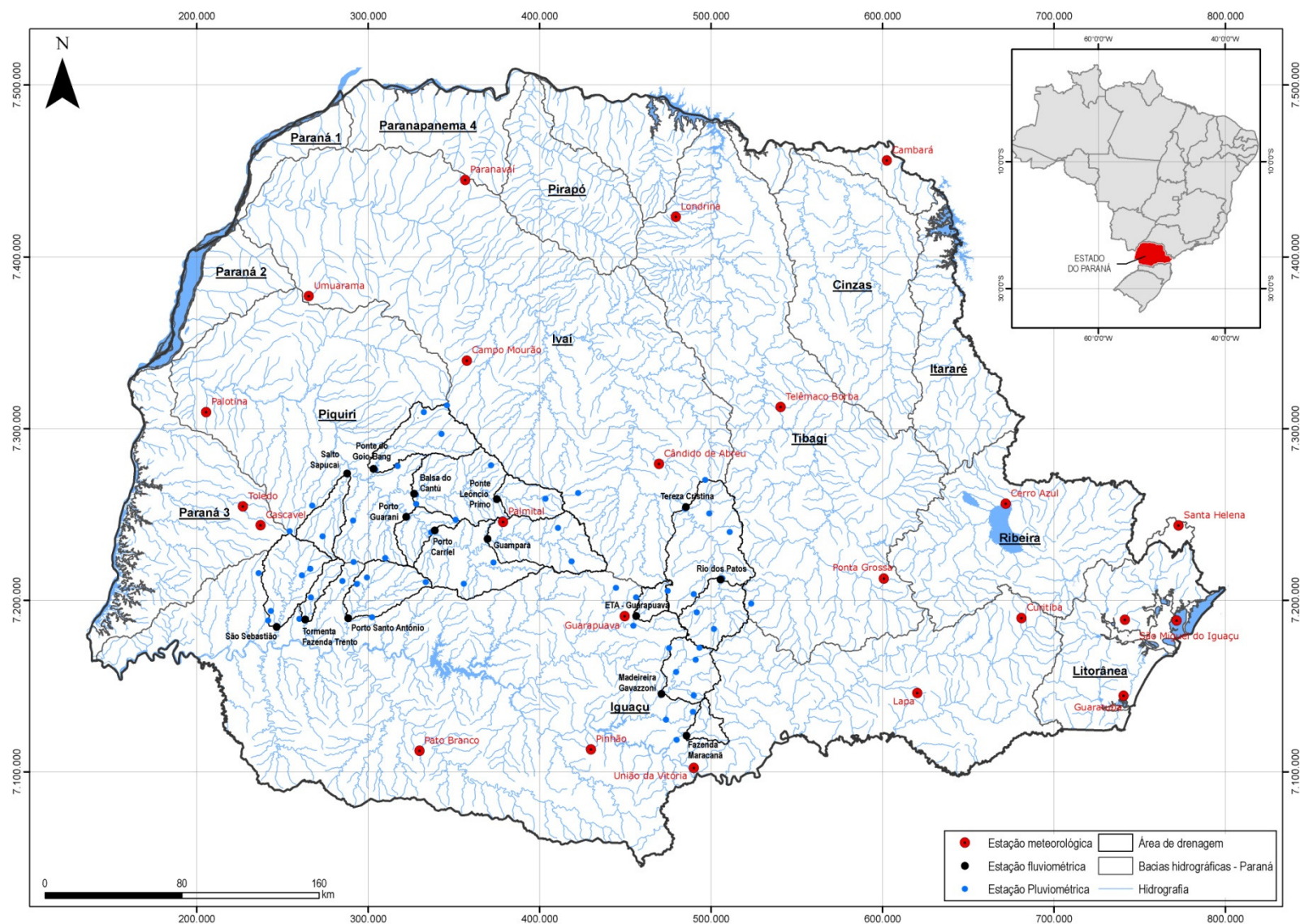


FIGURA 11: ESTAÇÕES HIDROMETEOROLÓGICAS UTILIZADAS NO ESTUDO.

FONTE: O AUTOR (2014)

TABELA 4: ESTAÇÕES METEOROLÓGICAS UTILIZADAS NO ESTUDO.

Município	Altitude (m)	Início de operação	Latitude (°)	Longitude (°)
Cambará	450	01/06/1997	-23,00	-50,00
Campo Mourão	601	02/06/1997	-24,05	-52,40
Cândido de Abreu	645	03/06/1997	-24,60	-51,30
Cascavel	719	04/06/1997	-24,90	-53,60
Cerro Azul	366	05/06/1997	-24,80	-49,30
Curitiba	935	06/06/1997	-25,40	-49,20
Pinhão	728	07/06/1997	-26,10	-51,70
Guarapuava	1070	10/06/1997	-25,40	-48,60
Guaratuba	0	11/06/1997	-25,40	-51,50
Lapa	909	12/06/1997	-25,80	-48,60
Londrina	585	13/06/1997	-25,80	-49,80
Palmital	783	14/06/1997	-23,30	-51,20
Palotina	303	15/06/1997	-24,90	-52,20
Paranavaí	480	16/06/1997	-24,30	-53,90
Pato Branco	721	17/06/1997	-23,10	-52,40
Ponta Grossa	885	18/06/1997	-26,10	-52,70
Santa Helena	271	19/06/1997	-25,20	-50,00
São Miguel do Iguaçu	298	20/06/1997	-24,90	-48,30
Telêmaco Borba	768	21/06/1997	-25,40	-48,30
Toledo	516	22/06/1997	-24,30	-50,60
Umuarama	480	23/06/1997	-24,80	-53,70
União da Vitória	756	24/06/1997	-23,70	-53,30

FONTE: O AUTOR ADAPTADO DE SIMEPAR (2012).

#### 4.2.1 Climatologia

Para as estações meteorológicas, os valores preenchidos dos pequenos períodos foram tomados em escala diária, uma vez que as estações apresentavam apenas poucos dias de falhas durante o período de 1998 a 2011. As falhas diárias foram preenchidas pelos dados brutos da estação mais próxima e prosseguiu-se com o cálculo mensal de suas variáveis temperatura, temperatura máxima, temperatura mínima, umidade relativa e radiação solar.

#### 4.2.2 Fluviometria

O preenchimento dos meses faltantes da série histórica de vazão das estações fluviométricas foi realizado a partir de uma regressão linear. A estação utilizada no preenchimento foi selecionada após verificada a semelhança hidrológica de escoamento superficial médio mensal com o posto a ser

preenchido, foi tomado como base um coeficiente de determinação ( $R^2$ ) superior a 0,80.

Cabe citar que o critério utilizado para definir se o mês apresenta ou não falhas está baseado na existência de, pelo menos, um único dia sem medições. Apesar de ser uma opção bem criteriosa, ela é válida para bacias hidrográficas de pequeno porte, uma vez que a variação diária de vazão pode implicar em médias mensais bem diferentes no caso de uma simples não consideração do dado faltante e da obtenção da média desconsiderando o mesmo.

#### 4.2.3 Pluviometria

No caso da precipitação, os meses com falhas das estações foram preenchidos a partir da Equação 27. Esse método se baseia na proporção linear entre postos vizinhos e é utilizado para dados mensais de precipitação, visando a homogeneização do período de informação das séries de precipitações.

Para um grupo de postos, são selecionados pelo menos três que possuam, no mínimo, 10 anos de dados. Para um posto  $Y$  que apresenta falhas, as mesmas são preenchidas com base na seguinte equação (Tucci, 2002):

$$y = \frac{1}{3} \left[ \frac{x_1}{x_{m1}} + \frac{x_2}{x_{m2}} + \frac{x_3}{x_{m3}} \right] \cdot Y_m \quad (27)$$

Onde  $y$  é a precipitação no posto  $Y$  no intervalo de tempo com falha,  $x_1$ ,  $x_2$  e  $x_3$  são as precipitações dos postos vizinhos no mesmo intervalo de tempo,  $x_{m1}$ ,  $x_{m2}$  e  $x_{m3}$  são as precipitações médias dos postos vizinhos e  $y_m$  é a precipitação média do posto onde será preenchido.

### 4.3 ESTRUTURAÇÃO DAS REDES NEURAIS ARTIFICIAIS (RNAs)

Ajustar uma RNA à solução de um problema nada mais é do que ajustar uma função a uma amostra de dados. No caso de sistemas ambientais complexos e dinâmicos, como é o caso da hidrologia, inexiste um tipo de função linear e pré definida que explique os fenômenos que ocorrem no sistema.

Portanto, o que se torna necessário é uma investigação das diversas maneiras de abordagem do problema.

Dentro da arquitetura de uma RNA vários são os parâmetros que podem ser alterados com a finalidade de obter um melhor ajuste do problema. Segundo Flood e Kartam (1994) *apud*. Sousa e Sousa (2006), é essencial investigar aspectos diferentes na modelagem com RNA a fim de melhorar a precisão da previsão dos processos hidrológicos. Os aspectos que podem ser alterados são: arquitetura da rede, processo de modelagem, forma de entrada dos dados e métodos de estimativa da eficiência para validação do modelo.

Nos estudos de Machado (2005) e Machado *et. al.* (2011), foi proposta uma metodologia de variação dos parâmetros das RNAs e recomendadas novas alterações. No presente estudo tais recomendações foram aceitas e algumas variações propostas foram aprimoradas.

Neste trabalho, as RNAs utilizadas também foram do tipo MLP com três camadas, sendo uma de entrada, uma intermediária e uma de saída. Todos os neurônios da camada intermediária possuem uma função de transferência do tipo sigmóide e a função de transferência da camada de saída é a função linear.

Com respeito aos modelos de entrada e saída, Machado (2005) e Machado *et. al.* (2011) testaram os mostrados na Tabela 1 e concluíram que aqueles que utilizaram como entrada dados de vazão e precipitação dos meses anteriores ( $t - 1$ ,  $t - 2$ ) obtiveram melhores resultados.

Para este trabalho, nem todas as conclusões de Machado (2005) e Machado *et. al.* (2011) foram agregadas, pois o objetivo do mesmo é a previsão de vazão em bacias hidrográficas sem monitoramento fluviométrico e, portanto, sem dados de vazão dos meses anteriores. Sendo assim, as recomendações foram utilizadas apenas para a precipitação e foram propostos os modelos de entrada apresentados na Tabela 5.

O número de neurônios criados na camada intermediária foi variado entre 3, 5, 8, 10, 15 e 30, enquanto que Machado (2005) e Machado *et. al.* (2011) variaram apenas entre 3, 5, 8 e 10.

Para as etapas de treinamento das RNAs, todas foram treinadas utilizando a técnica de otimização de Levenberg-Maquardt, com seus parâmetros  $\mu$  e  $\beta$  valendo respectivamente 0,1 e 10.



TABELA 5: ENTRADAS PROPOSTAS PELO AUTOR (2014).

Modelo	Entradas	Saídas
1	$Scores da ACP(P, T, T_{máx}, T_{mín}, U_{rel}, R_{solar})$	$Q(t)$
2	$P(t), P(t - 1), P(t - 2)$	$Q(t)$
3	$P(t), P(t - 1), P(t - 2), T(t)$	$Q(t)$
4	$P(t), P(t - 1), P(t - 2), R(t)$	$Q(t)$

ACP – ANÁLISE DE COMPONENTES PRINCIPAIS, P – PRECIPITAÇÃO MÉDIA MENSAL (mm/mês), T – TEMPERATURA (°C) , U – UMIDADE (%), R – RADIAÇÃO SOLAR (W/m²)

FONTE: O AUTOR (2014).

Assim como nos estudos de Machado (2005), optou-se por variar a extensão dos dados de entrada para calibração e validação das RNAs a fim de melhorar seu treinamento e, conseqüentemente, seus resultados. Como o número de dados disponíveis era referente a 168 meses (1998 a 2011), estes foram separados em 3 conjuntos, variando entre 60, 90 e 120 os dados para treinamento e 108, 78 e 48 os dados para validação, respectivamente.

Antes do início do treinamento, quatro foram os métodos de normalização utilizados para os conjunto de dados de entrada. Estes que foram sugeridos por Sajikumar e Thandaveswara (1999), Lima e Ferreira Filho (2003) , Rajurkar *et. al* (2002,2004) e Sperb *et. al.* (1999). Machado (2005) utilizou apenas a normalização proposta por Lima e Ferreira Filho (2003).

O número de *epochs* também foi variado, todas as RNAs foram inicializadas e treinadas com 30, 60, 90 e 120 *epochs*, seguindo a sugestão de Machado (2005). Além disso, todas as RNAs foram reinicializadas 6 vezes de modo que o peso inicial de seus neurônios fossem regenerados para cada nova inicialização.

Dessa maneira, cada resposta é o resultado de uma combinação diferente entre o número de entradas da RNA para calibração (60, 90 ou 120), o tipo de normalização (1 - Sajikumar e Thandaveswara (1999), 2 - Lima e Ferreira Filho (2003), 3 - Rajurkar *et. al* (2002,2004) ou 4 - Sperb *et. al.* (1999)), a inicialização (1, 2, 3, 4, 5, ou 6), o número de neurônios na camada intermediária (3, 5, 8, 10, 15 ou 30) e o número de *epochs* (10, 30, 60 ou 120). Estas combinações foram utilizadas na nomeação das RNAs, por exemplo, a RE90N4I5N8E30 se refere à RNA com 90 dados de entrada, normalização do tipo 4, 5 inicializações, 8 neurônios na camada intermediária e 30 *epochs*.

Ao final, para avaliar a influência dos elementos propostos no treinamento da RNA e na validação dos dados, as estatísticas consideradas foram

o coeficiente de determinação ( $R^2$ ), o erro médio quadrático (RMSE) e o coeficiente de Nash-Sutcliffe:

$$R^2 = \frac{[\sum_{t=1}^T (Q_o - \overline{Q_o}) \cdot (Q_m - \overline{Q_m})]^2}{\sum_{t=1}^T (Q_o - \overline{Q_o})^2 \cdot \sum_{t=1}^T (Q_m - \overline{Q_m})^2} \quad (28)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^T (Q_o - Q_m)^2}{n}} \quad (29)$$

$$NASH \ SUTCLIFFE = 1 - \frac{\sum_{t=1}^T (Q_o^t - Q_m^t)^2}{\sum_{t=1}^T (Q_o^t - \overline{Q_o})^2} \quad (30)$$

onde  $T$  é o número total de dados,  $Q_o$  é o valor de vazão observado,  $Q_m$  o valor de vazão modelado. O símbolo “ $\overline{\phantom{x}}$ ” refere-se ao valor médio.

#### 4.4 VALIDAÇÃO DO MÉTODO DE REGIONALIZAÇÃO

Para a verificação da validade do método proposto nesse trabalho, as estações Madeireira Gavazzoni e Guampará foram tomadas como bacias teste, ou seja, a partir de seus parâmetros físicos e fisiográficos, elas são sujeitas à análise discriminante para verificar em qual dos grupos formados pelas outras 13 estações elas se inserem e, depois disso, suas séries de vazões são geradas utilizando as RNAs calibradas para cada uma das estações contidas nos grupos nos quais elas foram alocadas.

Como resultado, é gerado um número de séries de vazões igual ao número de estações presentes no grupo. A formação de uma única série é realizada a partir do método do inverso do quadrado da distância, ou seja, ponderando os valores de vazão pelo inverso do quadrado da distância da estação base com as outras do mesmo grupo.

## 5 RESULTADOS

Os resultados apresentados nesse capítulo estão divididos em três partes: consistência dos dados hidrológicos, modelagem chuva-vazão e análise estatística multivariada.

### 5.1 CONSISTÊNCIA DOS DADOS FLUVIOMÉTRICOS

No tocante à análise de consistência dos dados hidrológicos, serão mostrados alguns resultados para ficar reforçada a importância do tratamento dos mesmos, haja vista a quantidade de erros nos dados disponibilizados pelo órgão competente.

A Figura 12 mostra os levantamentos de cota e vazão realizados em campo, a curva-chave estabelecida pela ANA e sua regeneração - fruto desse trabalho da estação ETA – Guarapuava no rio das Pedras. Observa-se que as medições mais altas de cota e vazão levantadas em campo (indicada pela seta ↑) não foram utilizadas no traçado da curva-chave pela ANA. Pelo gráfico, fica mais clara a importância que tais valores têm na geração da curva-chave, uma vez que, ao considerá-los no ajuste da curva, a mesma muda completamente o sentido de ascensão trazendo diferenças consideráveis na vazão correspondente aos valores das cotas mais altas. Verifica-se ainda que o método de Stevens trouxe pouca influência nos valores de vazão para as cotas mais altas, quando comparado com a extrapolação logarítmica.

No caso da estação Rio dos Patos, os dados de vazão disponibilizados são completamente inconsistentes e a sua utilização traria erros consideráveis na regionalização. A Figura 13 apresenta esses erros.

Verifica-se que a curva-chave sugerida pela ANA dispõe apenas de vazão nula para os anos posteriores a 1997, destoando, obviamente, da realidade do curso d'água em questão.

Além dos dois erros supracitados, outro posto que trouxe problemas na curva-chave foi a estação Ponte Leônicio Primo. Ao recalcular suas curvas-chave e avaliar o hidrograma anual médio das vazões novas obtidas e daquelas consistidas pela ANA foi possível verificar notável diferença a partir do ano de 1988 (Figura 14).

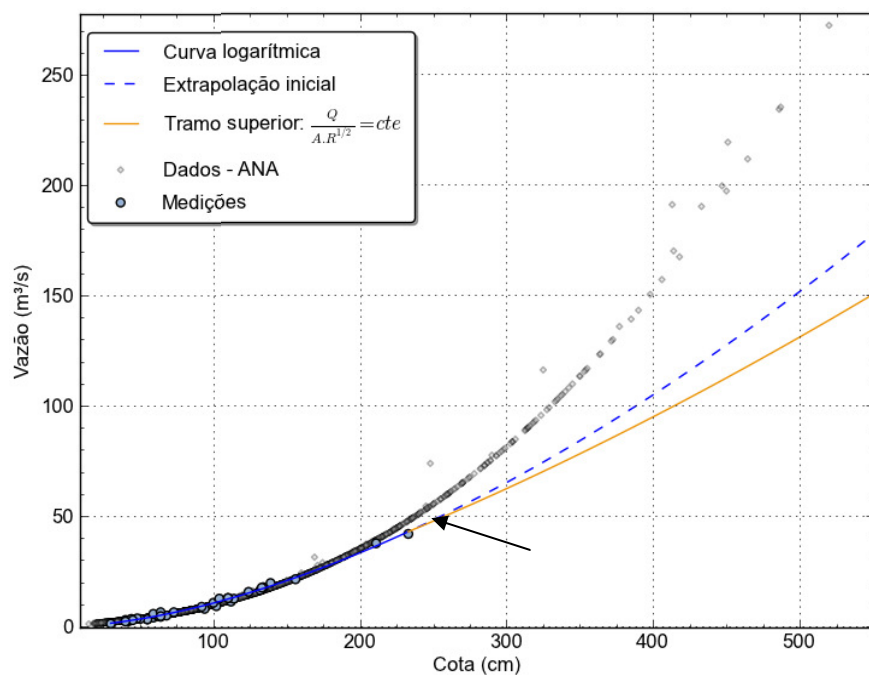


FIGURA 12: CURVA DE DESCARGA OBTIDA PARA A ESTAÇÃO ETA – GUARAPUAVA EM SEU 3º PERÍODO DE RÉGUA (17/05/1995 A 31/12/2012).

FONTE: O AUTOR (2014).

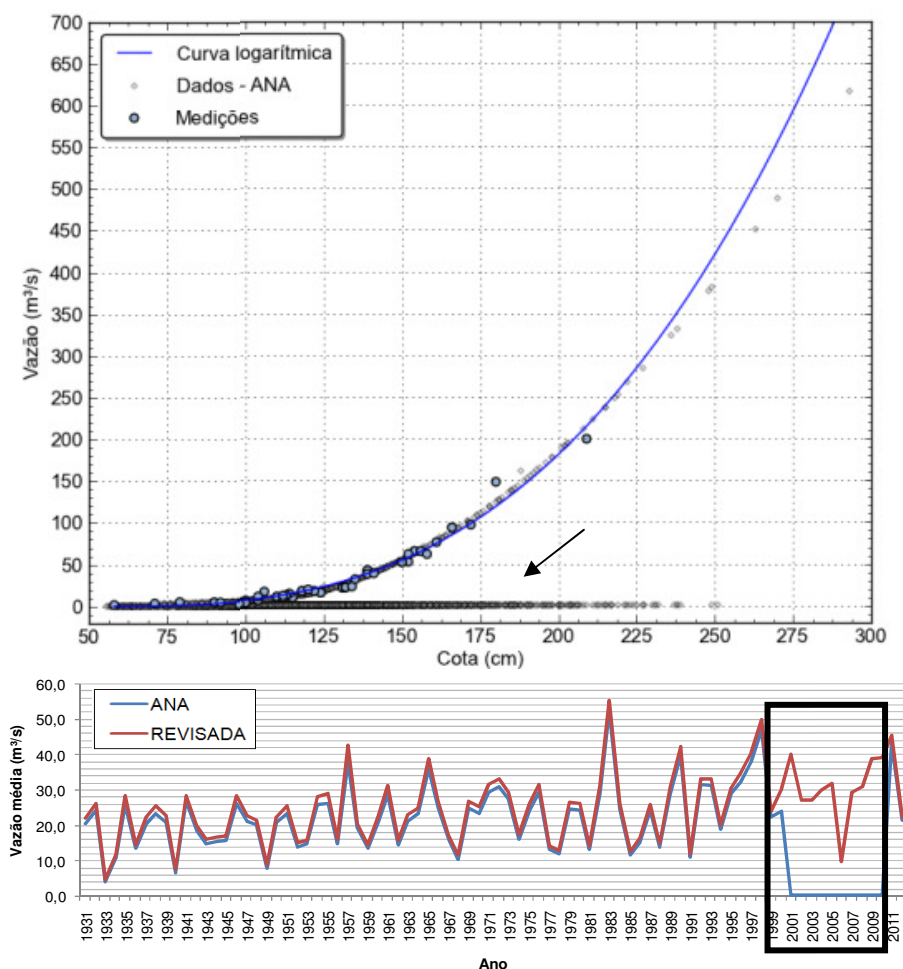


FIGURA 13: CURVA DE DESCARGA OBTIDA PARA A ESTAÇÃO RIO DOS PATOS EM SEU 3º PERÍODO DE RÉGUA (01/11/2000 A 31/12/2012).

FONTE: O AUTOR (2014).

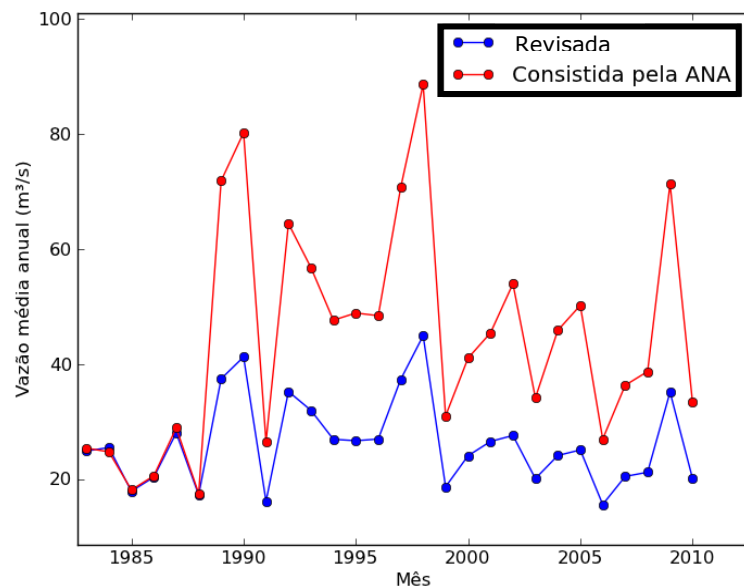


FIGURA 14: HIDROGRAMA ANUAL DA ESTAÇÃO PONTE LEÔNCIO PRIMO.  
FONTE: O AUTOR (2014).

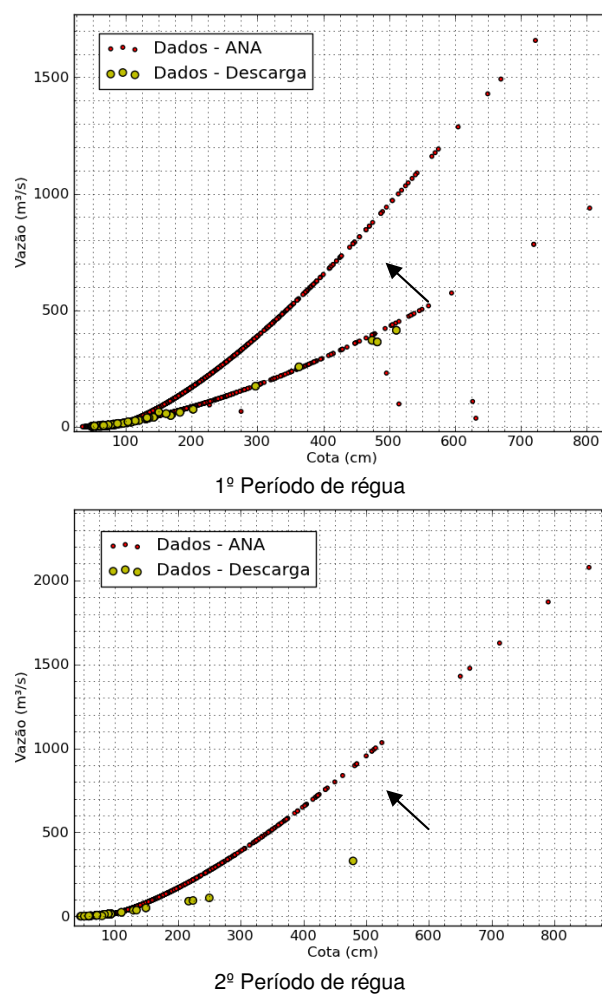


FIGURA 15: CURVA DE DESCARGA OBTIDA PARA A ESTAÇÃO PONTE LEÔNCIO PRIMO EM SEUS DOIS PERÍODOS DE RÉGUA DIFERENTES.

FONTE: O AUTOR (2014).

Analisando mais a fundo o problema tornou-se possível detectar a origem do erro. No momento do traçado da curva-chave por parte do Instituto das Águas do Paraná para o primeiro período de régua, houve algum erro por parte do órgão que a curva tomou um sentido completamente distorcido em relação aos levantamentos de campo. Os gráficos mostrados na Figura 15 ilustram o problema.

## 5.2 USO DA RNA PARA MODELAGEM CHUVA-VAZÃO

Com respeito às simulações chuva-vazão a partir de RNAs, um cuidado foi necessário ser tomado. Como o objetivo do trabalho é a regionalização de vazões a partir da utilização de RNAs calibradas e validadas em bacias hidrográficas com monitoramento pluviométrico e climatológico, todas as bacias devem ser modeladas a partir do mesmo número de dados de entrada. Por exemplo, não seria possível calibrar e validar uma RNA com dados de precipitação de 7 estações e regionalizar tal RNA para uma bacia hidrográfica onde não existe sequer 1 estação pluviométrica localizadas dentro dos limites de sua bacia, pois faltariam dados de entrada à ela.

Por isso, somado ao fato de que os melhores resultados foram provenientes da aplicação dos dados de precipitação e temperatura (Modelo de entrada 3 da Tabela 5), antes da calibração das RNAs, foi determinado que todas deveriam ser treinadas e validadas a partir de 4 estações pluviométricas e uma série mensal de dados de temperatura. Para a determinação de quais estações pluviométricas utilizar nas bacias que contam com mais de 4 foi realizada uma análise de componentes principais. A Tabela 6 mostra os resultados obtidos.

Na última coluna da tabela está apresentada a variância cumulativa das estações selecionadas para explicar o fenômeno da chuva nas áreas de drenagem. Verifica-se que o autor optou por selecionar variâncias superiores a 90%.

TABELA 6: RESULTADOS DA ANÁLISE DE COMPONENTES PRINCIPAIS PARA UTILIZAÇÃO DAS ESTAÇÕES PLUVIOMÉTRICAS COMO ENTRADA PARA AS RNAS.

Código	Estação Fluviométrica	Estações Pluviométricas	ACP	% da Variância
64625000	TEREZA CRISTINA	Apiaba, Bairro da Imbuia, Campo de Dentro, Inácio Martins, Itapara, Laranjeiras, Relógio, Rio dos Patos, Saltinho, Tereza Cristina e Tijuco Preto	Apiaba, Campo de Dentro, Itapara e Laranjeiras	93%
64620000	RIO DOS PATOS	Apiaba, Itapara, Relógio, Rio dos Patos, Tijuco Preto e Inácio Martins	Apiaba, Relógio, Tijuco Preto e Inácio Martins	97%
65809000	ETA - GUARAPUAVA	Guarapuava, Campo de Dentro, Invernadinha e Palmeirinhas	N/A	N/A
65764000	MADEIREIRA GAVAZZONI	Campina Bonita, Colônio Alegrave, Fazenda Zaniolo, Inácio Martins, Leonópolis e Madeira Gavazzoni	Campina Bonita, Fazenda Zaniolo, Inácio Martins e Leonópolis	97%
65415000	FAZENDA MARACANÃ	Fazenda Maracanã, Pátio Velho, Salto do Vau e Santana	N/A	N/A
65971050	TORMENDA FAZENDA TRENTO	Catanduvas, Flor da Serra, Ibema, Ibiracema, Nelson Taborda e Rio do Salto	Catanduvas, Flor da Serra, Ibema e Ibiracema	97%
65979000	SÃO SEBASTIÃO	Boi Preto, Cascavel – Ocepar, Flor da Serra, Linha Santa Catarina, Nelson Taborda, Rio do Salto, Santa Lúcia e São João do Oeste	Boi Preto, Cascavel – Ocepar, Nelson Taborda e Rio do Salto	96%
64775000	BALSA DO CANTÚ	Altamira do Paraná, Balsa do Cantú, Barra Grande, Laranjal, Palmital e Roncador	Balsa do Cantú, Barra Grande, Laranjal e Roncador	96%
64764000	GUAMPARÁ	Campina do Simão, Marquinhos, Palmital e Santa Maria do Oeste	N/A	N/A
64767000	PORTO CARRIEL	Campina do Simão, Campo Verde, Laranjal, Marquinhos, Palmital, Porto Carriel e Santa Maria do Oeste	Campina do Simão, Marquinhos, Palmital e Porto Carriel	97%
64773000	PONTE LEÔNCIO PRIMO	Barra Grande, Roncador, Palmital e Pitanga	N/A	N/A
64785000	PONTE DO GOIO BANG	Campina da Lagoa, Guarani, Mamborê, Pensamento e Ponte do Goio-Bang	Campina da Lagoa, Guarani, Mamborê e Ponte do Goio-Bang	99%
64790000	SALTO SAPUCAÍ	Corbélia, Ponte Tourinho – Braganey, Salto Sapucaí e São João do Oeste	N/A	N/A
65970001	PORTO SANTO ANTÔNIO	Goioxim, Guaraniaçu, Linha Mirim, Porto Santo Antônio, Torra Alta e Vila Guarani	Goioxim, Linha Mirim, Torra Alta e Vila Guarani	97%
64771500	PORTO GUARANI	Campina do Simão, Campo Verde, Laranjal, Marquinhos, Palmital, Porto Carriel e Santa Maria do Oeste	Campina do Simão, Marquinhos, Palmital e Porto Carriel	97%

FONTE: O AUTOR (2014).

Definidas as estações e geradas as séries históricas de temperatura para o centróide das bacias hidrográficas, as RNAs foram calibradas e validadas. A Tabela 7 mostra os resultados e estatísticas obtidas para as 15 estações fluviométricas analisadas e as figuras de 16, 17, e disponíveis no Apêndi-

ce 1, mostram os hidrogramas resultantes da calibração e validação do modelo.

Como observado por Machado (2002), não existe um padrão na estruturação para o melhor desempenho das RNAs. Ao passo que oito estações precisaram de 90 dados de entrada para a melhor calibração do modelo (ETA – Guarapuava, Fazenda Maracanã, Tormenta Fazenda Trento, São Sebastião, Porto Carriel, Ponte do Goio Bang, Porto Santo Antônio e Porto Guarani), três precisaram de 120 (Rio dos Patos, Balsa do Cantú e Salto Sapucaí) e duas de 60 (Tereza Cristina e Ponte Leôncio Primo). Não houve nenhuma correlação entre o número de dados de entrada e o tipo de normalização e o número de neurônios na camada intermediária. Com respeito ao número de *epochs*, seis das 13 RNAs deste trabalho tiveram melhor desempenho com apenas 30 reinicializações, ou seja, a hipótese levantada por Machado (2002) foi novamente comprovada – o algoritmo de treinamento de Levenberg Maquardt trata-se de uma técnica de convergência rápida. O número de reinicializações dos pesos e vieses também foram muito variados e sem correlação notável com os outros parâmetros variados.

Com respeito às vazões modeladas e observadas durante a validação dos modelos, os resultados foram bem satisfatórios. Avaliando pelo coeficiente de Nash Sutcliffe, a estação que obteve o melhor índice foi a Porto Carriel, com o valor de 0,87. O pior resultado foi o da estação Ponte do Goio Bang, com 0,67, porém, quando analisado em conjunto com o coeficiente de determinação (0,78) e o erro médio quadrático (6,75), seu resultado se destaca diante de várias outras estações.

Pelas vazões médias e desvios padrões avaliados, verificou-se que as maiores diferenças entre a série modelada e observada ocorrem nos desvios padrões. Enquanto que a diferença máxima, em módulo, nas vazões médias é de aproximadamente 15%, nos desvios padrões essa diferença chega a 31%, e ambas ocorrem para a estação Ponte do Goio Bang. A estação Fazenda Maracanã foi a que apresentou, quando somados, os menores resultados, a diferença das vazões médias observadas e modeladas durante a validação resultou em 4% e o desvio padrão em 3%.



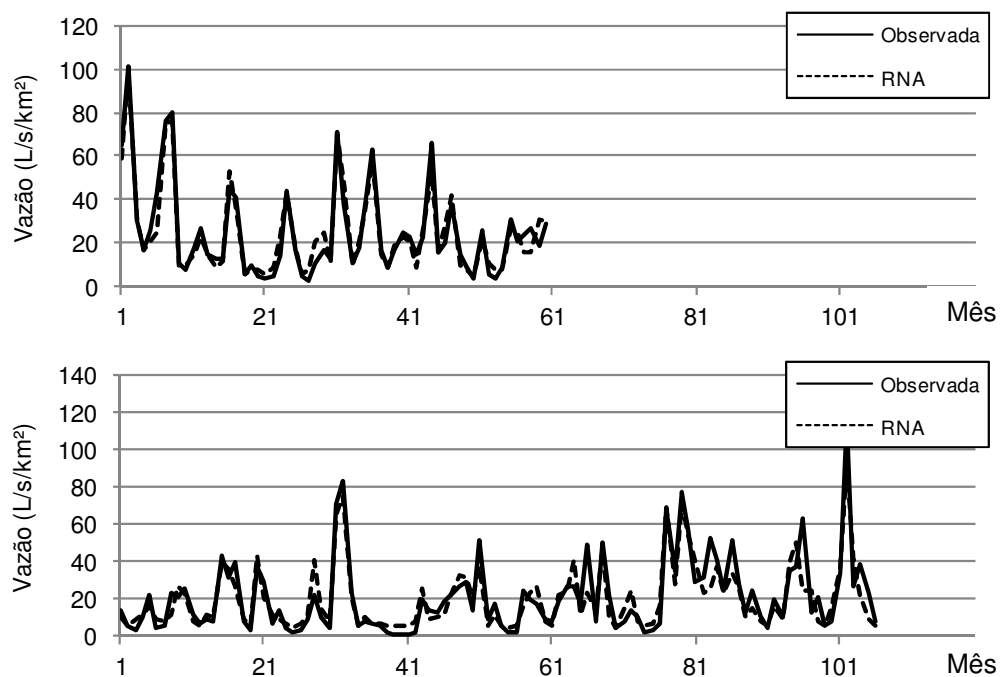


FIGURA 16: CALIBRAÇÃO E VALIDAÇÃO DAS RNAs PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO TEREZA CRISTINA.

FONTE: O AUTOR (2014).

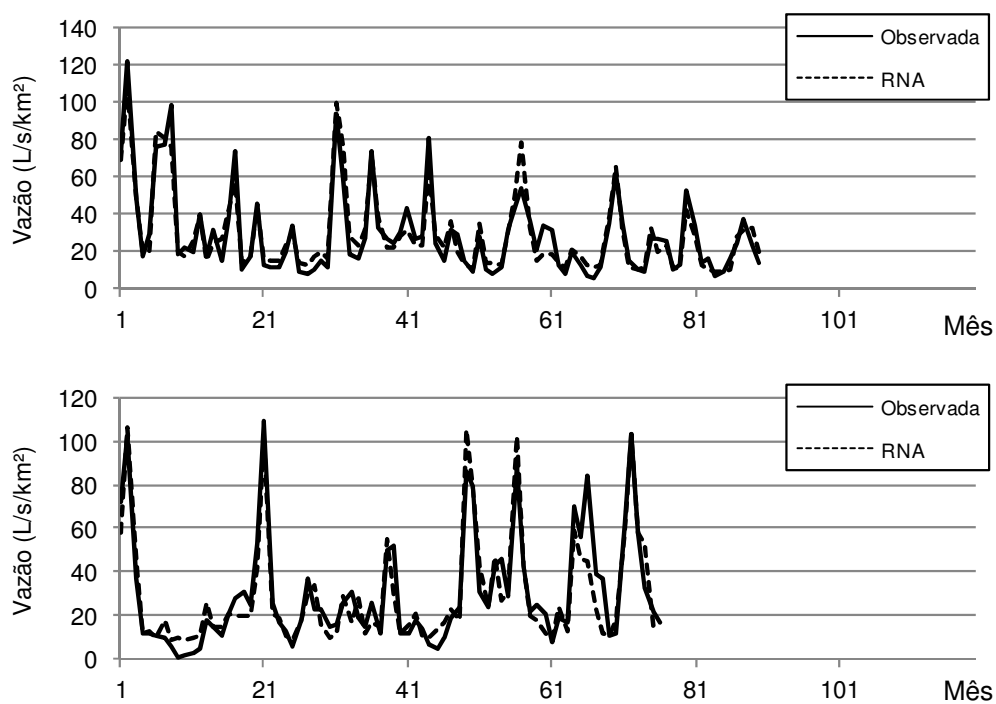


FIGURA 17: CALIBRAÇÃO E VALIDAÇÃO DAS RNAs PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO FAZENDA MARACANÃ.

FONTE: O AUTOR (2014).

TABELA 7: RESULTADOS OBTIDOS DA APLICAÇÃO DAS RNAs PARA MODELAGEM CHUVA-VAZÃO DAS BACIAS HIDROGRÁFICAS.

ESTAÇÃO	REDE NEURAL	R <sup>2</sup>		RMSE		NASH SUT-CLIFFE		CALIBRAÇÃO				VALIDAÇÃO			
		C.	V.	C.	V.	C.	V.	Q <sub>OBS</sub> (l/s/km <sup>2</sup> )	Q <sub>RNA</sub> (l/s/km <sup>2</sup> )	S <sub>OBS</sub> (l/s/km <sup>2</sup> )	S <sub>RNA</sub> (l/s/km <sup>2</sup> )	Q <sub>OBS</sub> (L/s/km <sup>2</sup> )	Q <sub>RNA</sub> (L/s/km <sup>2</sup> )	S <sub>OBS</sub> (L/s/km <sup>2</sup> )	S <sub>RNA</sub> (L/s/km <sup>2</sup> )
TEREZA CRISTINA	RE60N2I6N10E90	0,93	0,81	5,54	8,83	0,93	0,81	25,49	25,37	21,66	20,40	21,69	20,49	20,16	17,35
RIO DOS PATOS	RE120N3I5N15E30	0,93	0,85	5,98	12,36	0,93	0,82	27,17	27,08	23,17	22,05	35,86	32,71	29,83	24,11
ETA - GUARAPUAVA	RE90N2I4N3E30	0,93	0,83	5,83	10,09	0,93	0,81	28,57	28,55	21,71	20,49	30,84	27,66	23,03	20,07
FAZENDA MARACANÃ	RE90N4I1N8E60	0,90	0,87	7,21	9,49	0,90	0,86	29,91	29,78	23,05	21,69	30,83	29,67	25,75	25,02
TORMENDA FAZENDA TRENTO	RE90N3I3N15E60	0,84	0,82	11,76	11,26	0,84	0,78	35,32	35,34	29,32	25,56	25,57	28,62	24,39	18,86
SÃO SEBASTIÃO	RE90N1I2N8E30	0,89	0,78	8,54	11,97	0,89	0,76	33,37	33,36	25,55	24,16	25,47	28,90	24,47	19,59
BALSA DO CANTÚ	RE120N1I2N5E30	0,88	0,89	7,96	9,10	0,88	0,85	27,02	26,85	23,25	21,09	30,77	26,26	23,89	20,74
PORTO CARRIEL	RE90N2I3N3E60	0,89	0,88	8,69	9,47	0,88	0,87	29,68	28,81	25,69	23,19	27,70	30,14	26,68	23,48
PONTE LEÔNCIO PRIMO	RE60N3I2N30E30	0,80	0,70	16,25	18,82	0,80	0,69	40,89	40,83	36,34	30,62	33,37	34,99	34,08	31,11
PONTE DO GOIO BANG	RE90N4I1N8E90	0,83	0,78	4,37	6,75	0,83	0,67	21,13	21,13	10,80	9,82	21,48	18,33	11,88	8,15
SALTO SAPUCAÍ	RE120N3I5N5E60	0,90	0,77	5,68	7,93	0,90	0,74	26,97	26,97	17,71	16,67	27,67	25,22	15,61	12,47
PORTO SANTO ANTÔNIO	RE90N1I6N15E30	0,91	0,82	8,71	9,65	0,91	0,82	29,64	28,05	28,95	26,70	23,41	23,04	23,12	21,52
PORTO GUARANI	RE90N1I2N10E90	0,89	0,87	8,27	9,59	0,89	0,86	28,95	27,54	25,10	22,84	28,03	29,07	26,09	23,24

C = CALIBRAÇÃO, V = VALIDAÇÃO.

FONTE: O AUTOR (2014).

### 5.3 REGIONALIZAÇÃO HIDROLÓGICA

Foram levantados para cada uma das 15 estações fluviométricas e de sua respectiva área de drenagem os seguintes parâmetros físicos: temperatura média, coeficiente de variação da temperatura média, precipitação anual média e coeficiente de variação da precipitação anual média, além dos parâmetros fisiográficos como altitude do ponto de instalação da estação fluviométrica, área de drenagem, densidade de drenagem, declividade média, altitude, latitude e longitude do centróide da bacia, totalizando 11 variáveis.

Antes da aplicação da análise de agrupamentos nas 13 bacias avaliadas (duas foram tomadas como bacias teste), foi necessária a realização de uma análise de componentes principais na amostra, isto a fim de eliminar dados que são simplesmente combinações lineares de outros dados e que, estatisticamente, não agregavam valor à amostra.

Após sua aplicação, pode-se perceber que apenas cinco variáveis (temperatura média, área de drenagem, precipitação média anual e longitude do centróide da bacia) já explicavam 95% da variância da amostra. Os resultados estão mostrados na Tabela 8, na qual a última coluna representa a porcentagem de variância acumulada que cada um dos parâmetros físico e fisiográficos agregam à amostra.

TABELA 8: RESULTADO DA ANÁLISE DE COMPONENTES PRINCIPAIS DOS PARÂMETROS LEVANTADOS PARA AS ESTAÇÕES FLUVIOMÉTRICAS ESTUDADAS.

Variável	Parâmetro Físico/Fisiográfico	% da Variância
1	Temperatura Média	57%
2	Área de drenagem	76%
3	Precipitação média anual	86%
4	Longitude do centróide da área de drenagem	91%
5	Coeficiente de variação da precipitação	95%
6	Declividade média da área de drenagem	97%
7	Altitude do centróide da área de drenagem	99%
8	Latitude do centróide da área de drenagem	99%
9	Altitude da estação fluviométrica	100%
10	Coeficiente de variação da temperatura	100%
11	Densidade de drenagem	100%

FONTE: O AUTOR (2014).

A análise de agrupamentos foi realizada considerando a formação de diversos grupos com o número de variáveis variando entre quatro e cinco, como ordenadas na Tabela 8.

Verificou-se primeiramente que, independente do número de variáveis utilizadas no agrupamento, a formação de três grupos resultava no valor de silhueta mais próximo da unidade, indicando, assim, que os indivíduos contidos em cada um dos grupos estavam mais próximos entre si e mais distantes dos outros grupos (Figura 18).

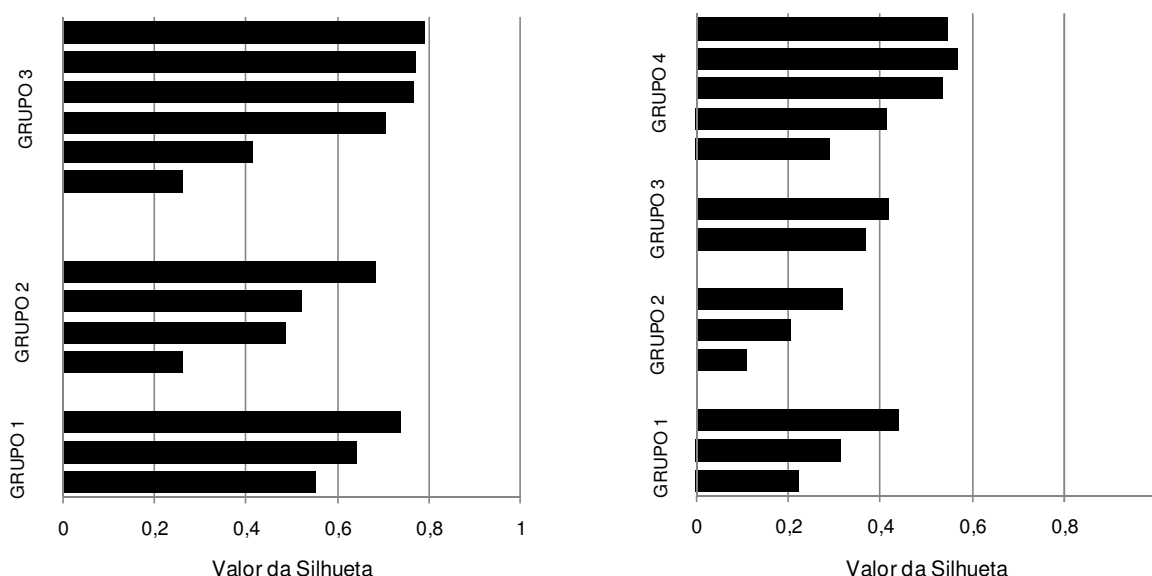


FIGURA 18: GRÁFICO DA SILHUETA PARA TRÊS (ESQUERDA) E QUATRO (DIREITA) GRUPOS FORMADOS.

FONTE: O AUTOR (2014).

Posteriormente, ao avaliar a formação dos três grupos a partir de quatro ou cinco variáveis, os mesmos indivíduos compunham cada um dos grupos. Ou seja, as informações adicionadas não estavam agregando informação à amostra e, dessa maneira, optou-se por utilizar apenas quatro variáveis para caracterização de cada estação fluviométrica analisada.

A Tabela 9 mostra uma síntese das variáveis utilizadas na análise de agrupamentos. Em grifado estão mostradas as bacias teste selecionadas.

Cabe citar que antes dos dados serem agrupados, eles foram devidamente normalizados em função do desvio padrão, de modo que no momento do cálculo das distâncias entre os indivíduos, unidades diferentes de escala

não trouxessem vantagens dos maiores em relação aos menores. Os grupos finais definidos estão mostrados na Tabela 10.

TABELA 9: PARÂMETROS FÍSICOS E FIOGRÁFICOS OBTIDOS PARA AS ÁREAS DE DRENAGEM DAS ESTAÇÕES FLUVIOMÉTRICAS ANALISADAS.

Código	Nome	Área de drenagem (km <sup>2</sup> )	Temperatura média (°C)	Precipitação anual (mm)	Longitude (m)
64775000	BALSA DO CANTÚ	2520,1	20,2	1746	7.264.100
65809000	ETA - GUARAPUAVA	312,8	18,1	1814	7.199.190
65415000	FAZENDA MARACANÃ	323,6	18,2	1617	7.126.860
<u>64764000</u>	<u>GUAMPARÁ</u>	<u>1688,7</u>	<u>19,9</u>	<u>1893</u>	<u>7.233.170</u>
<u>65764000</u>	<u>MADEIREIRA GAVAZZONI</u>	<u>935,7</u>	<u>18,5</u>	<u>1884</u>	<u>7.158.460</u>
64785000	PONTE DO GOIO BANG	1341,8	20,6	1818	7.294.420
64773000	PONTE LEÔNIO PRIMO	696,6	20,1	1997	7.258.480
64767000	PORTO CARRIEL	3569,3	20,1	1946	7.227.640
64771500	PORTO GUARANI	4165	20,1	1947	7.229.420
65970001	PORTO SANTO ANTÔNIO	1084,1	20,2	1995	7.206.870
64620000	RIO DOS PATOS	1097,7	18,8	1780	7.191.890
64790000	SALTO SAPUCAÍ	693,3	20,6	1906	7.246.730
65979000	SÃO SEBASTIÃO	1311,6	20,6	1963	7.213.920
64625000	TEREZA CRISTINA	3567,9	19,1	1729	7.219.170
65971050	TORMENTA FAZENDA TRENTA	319,6	20,5	1975	7.207.560

FONTE: O AUTOR (2014).

TABELA 10: GRUPOS DE ESTAÇÕES FLUVIOMÉTRICAS FORMADOS.

Grupo 1	Grupo 2	Grupo 3
PONTE DO GOIO BANG	BALSA DO CANTÚ	ETA - GUARAPUAVA
PONTE LEÔNIO PRIMO	PORTO CARRIEL	FAZENDA MARACANÃ
PORTO SANTO ANTÔNIO	PORTO GUARANI	RIO DOS PATOS
SALTO SAPUCAÍ	TEREZA CRISTINA	
SÃO SEBASTIÃO		
TORMENTA FAZENDA TRENTA		

FONTE: O AUTOR (2014).

TABELA 11: RESULTADO DA ANÁLISE DISCRIMINANTE APLICADA ÀS BACIAS TESTE.

Grupo	Estação Fluviométrica	Distância
Grupo 1	Guampará	2,633454996
Grupo 2	Guampará	1,751589888
Grupo 3	Guampará	4,219637069
Grupo 1	Madeira Gavazzoni	3,650189888
Grupo 2	Madeira Gavazzoni	4,284854996
Grupo 3	Madeira Gavazzoni	2,852437069

FONTE: O AUTOR (2014).

Depois de realizada a análise discriminante, verificou-se que a estação Guampará encontra-se mais próxima do centróide do Grupo 2, enquanto que a estação Madeireira Gavazzoni encontra-se mais próxima do Grupo 3. Na Tabela 11 estão mostradas as distâncias obtidas de cada um.

Sendo assim, a série de vazões mensais regionalizada para a estação Guampará foi regenerada a partir das RNAs calibradas das estações componentes do Grupo 2: Balsa do Cantú, Porto Carriel, Porto Guarani e Tereza Cristina e a série da estação Madeireira Gavazzoni a partir das RNAs das estações do Grupo 3: ETA – Guarapuava, Fazenda Maracanã e Rio dos Patos.

Após aplicação das RNAs calibradas a partir dos dados de precipitação e temperatura de cada uma das estações, foram obtidas quatro e três séries de vazões mensais para as estações Guampará e Madeireira Gavazzoni, respectivamente. De modo a unificar essas séries, às mesmas foi aplicado uma ponderação a partir do método do inverso do quadrado da distância, gerando apenas uma série representativa das estações teste.

Na Tabela 12 e Tabela 13 estão mostradas a distância entre as estações do mesmo grupo e as estações base e algumas estatísticas das séries de vazões calculadas a partir das RNAs. Na penúltima coluna, em negrito, estão mostradas as estatísticas da vazão regionalizada para as estações base a partir do cálculo do inverso do quadrado da distância das outras estações do mesmo grupo. E na última coluna estão apresentadas a média e o desvio padrão das séries de vazões observadas nas estações base.

TABELA 12: DISTÂNCIAS E PRINCIPAIS ESTATÍSTICAS DA SÉRIE REGIONALIZADA PARA A ESTAÇÃO GUAMPARÁ.

Estação	Balsa do Cantú	Porto Carriel	Porto Guarani	Tereza Cristina	<b>Regionalizada para estação Guampará</b>	<b>Guampará</b>
Distância (km)	50,5	31,6	48,7	116,9	-	-
Média (L/s/km <sup>2</sup> )	25,07	29,63	28,34	21,75	<b>28,09</b>	<b>30,20</b>
Desvio Padrão (L/s/km <sup>2</sup> )	20,94	23,39	23,00	18,13	<b>22,25</b>	<b>24,45</b>
Nash-Sutcliffe	0,73	0,80	0,80	0,57	<b>0,82</b>	-
R <sup>2</sup>	0,78	0,81	0,81	0,63	<b>0,82</b>	-

FONTE: O AUTOR (2014).

TABELA 13: DISTÂNCIAS E PRINCIPAIS ESTATÍSTICAS DA SÉRIE REGIONALIZADA PARA A ESTAÇÃO MADEIREIRA GAVAZZONI.

Estação	ETA - Guarapuava	Fazenda Maracanã	Rio dos Patos	<b>Regionalizada para estação Madeireira Gavazzoni</b>	<b>Madeiraira Gavazzoni</b>
Distância (km)	48,1	28,9	75,1	-	-
Média (L/s/km <sup>2</sup> )	23,54	32,68	32,61	<b>30,49</b>	<b>30,34</b>
Desvio Padrão (L/s/km <sup>2</sup> )	15,58	23,60	26,68	<b>21,34</b>	<b>23,36</b>
Nash-Sutcliffe	0,58	0,84	0,75	<b>0,88</b>	-
R <sup>2</sup>	0,70	0,86	0,82	<b>0,88</b>	-

FONTE: O AUTOR (2014).

Percebe-se que para ambas as estações os resultados obtidos foram satisfatórios. As diferenças obtidas para as médias foram de 7,5% na estação Guampará e 0,5% na estação Madeireira Gavazzoni, para os desvios foram de 9,9% para a estação Guampará e 9,5% para a estação Madeireira Gavazzoni.

As Figura 19 e Figura 20 mostram uma comparação entre as séries de vazões obtidas pelo método proposto de regionalização nesse trabalho e as vazões observadas nas estações teste.

Nota-se que os eventos extremos, tanto de cheia como de estiagem, foram bem representados, com pouca dispersão em torno da regressão linear, e os valores das estatísticas calculadas corresponderam às expectativas, uma vez que as médias e desvios se apresentaram bastante semelhantes e os coeficientes de Nash-Sutcliffe e de determinação bem elevados.

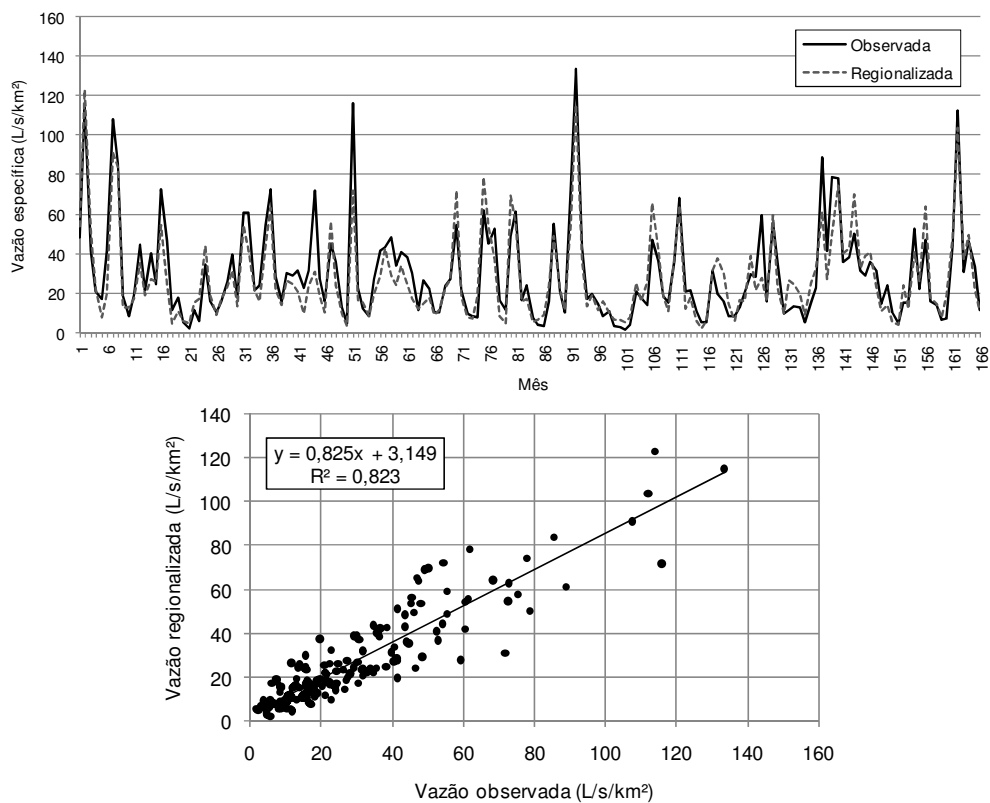


FIGURA 19: VAZÕES REGIONALIZADAS E OBSERVADAS – ESTAÇÃO GUAMPARÁ.  
FONTE: O AUTOR (2014).

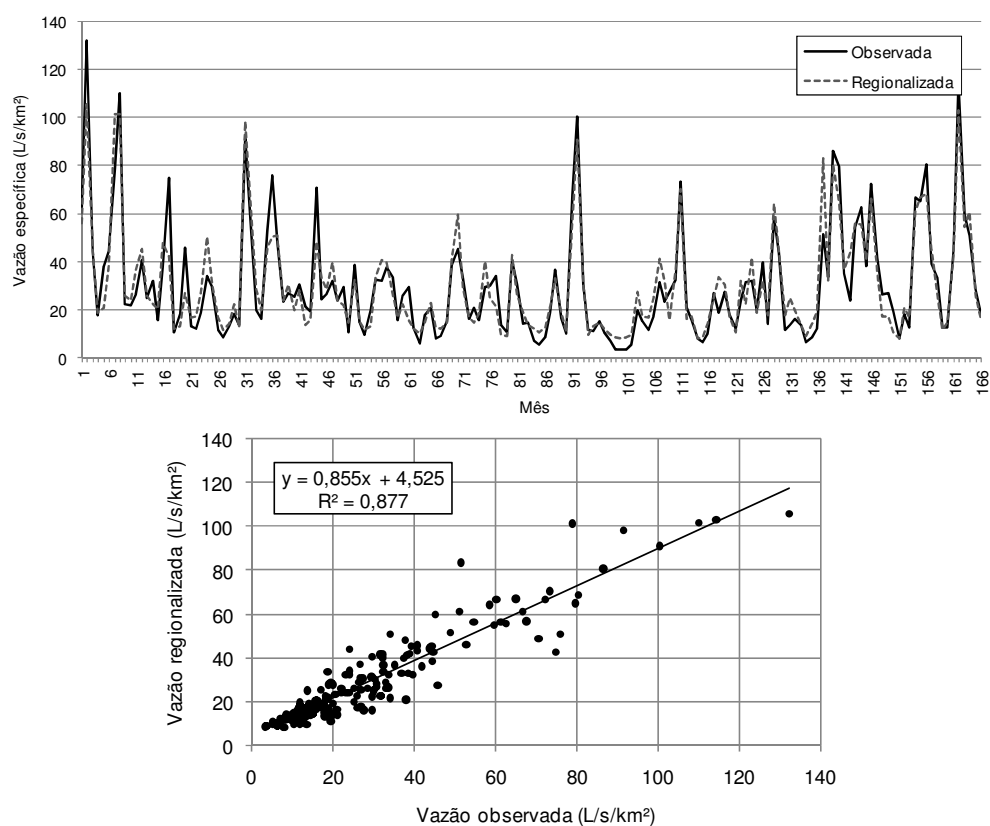


FIGURA 20: VAZÕES REGIONALIZADAS E OBSERVADAS – ESTAÇÃO MADEIREIRA GA-  
VAZZONI.  
FONTE: O AUTOR (2014).



## 6 CONCLUSÕES E RECOMENDAÇÕES

Pelo que foi apresentado nos capítulos 3, 4 e 5 dessa dissertação comprova-se que os objetivos gerais e específicos propostos foram atingidos. A seguir estão apresentadas as principais conclusões e recomendações para a realização de estudos futuros.

Pelos resultados apresentados, ficou clara a importância da consistência dos dados hidrológicos disponibilizados no sítio eletrônico da ANA (*Hidroweb*). Após a avaliação fluviométrica detalhada principalmente das estações ETA – Guarapuava, Ponte Leôncio Primo e Rio dos Patos a série histórica de vazões destas estações ficaram mais realistas e coerentes, isso devido às correções de traçado da curva-chave para a região onde existiam dados levantados em campo e às extrapolações estabelecidas por métodos já consagrados na literatura, como o de Stevens.

Por esse trabalho ter sido concebido com a finalidade gerar, a partir de uma técnica de regionalização, séries históricas de vazões para aproveitamentos hidrelétricos de pequeno porte, a extrapolação da curva-chave pelo método de Stevens nem sempre foi utilizada, haja vista que em alguns casos ela subestimava consideravelmente as vazões extremas, o que poderia incorrer em riscos às estruturas de segurança que seriam dimensionadas para os aproveitamentos.

Outro ponto importante na consistência dos dados fluviométricos foi a análise de informações dos períodos de régua das estações. Muitas vezes, alguns dados de cota e vazão levantados de campo estavam sendo utilizados para o traçado da curva-chave de períodos diferentes dos quais eles haviam sido levantados, acarretando em erros sistemáticos à série de vazões.

No geral, com relação à modelagem chuva-vazão por meio de RNAs, cabe citar que os resultados obtidos por este modelo empírico foram satisfatórios. Os índices de determinação, erro médio quadrático e coeficientes de Nash-Sutcliffe obtidos para todas as estações foram bons e dentro do que a literatura configura como confiável para estudos hidrológicos. Se tomados os

devidos cuidados no momento da calibração das RNAs, sua aplicação no ramo da hidrologia pode ser ainda muito desenvolvida.

Como foi constatado nesse trabalho, quanto maior a variância dos dados de entrada utilizados nas RNAs, melhores são os resultados. A ideia inicial desse trabalho era a utilização apenas da precipitação média mensal na bacia hidrográfica (obtida pelo método de Thiessen) como dado de entrada das RNAs já treinadas e calibradas. Porém, após alguns testes, verificou-se que os resultados obtidos por esse método não eram tão satisfatórios quanto quando utilizados como de entrada dados brutos das estações pluviométricas inseridas no entorno da bacia. Esta maior variância inserida na amostra trazia mais informações às redes, fazendo com que elas “aprendessem” melhor as características no entorno da região e levassem à melhores resultados no momento de validação.

Além disso, verificou-se também que as RNAs se aplicam melhor em locais onde a vazão é mais sensível à precipitação, ou seja, onde a resposta de um fenômeno de precipitação é dada de maneira mais rápida, como em bacias formadas por solos de textura argilosa, com menor taxa de infiltração, como na região da formação Serra Geral no Estado do Paraná, analisada nesse trabalho. Verificou-se isso pelos resultados ruins obtidos da calibração das RNAs em bacias hidrográficas inseridas na porção norte e noroeste do Estado, que compreende as regiões mais baixas dos rios Piquiri e Ivaí. Recomenda-se, portanto, a aplicação do método apresentado nessa dissertação apenas para as bacias hidrográficas incluídas nos limites geográficos das bacias hidrográficas aqui analisadas.

Os melhores resultados de adaptação das RNAs foram obtidos ao utilizar como dados de entrada a precipitação em três períodos diferentes ( $t$ ,  $t - 1$  e  $t - 2$ ) e a temperatura no tempo  $t$ . No trabalho, ficou clara a importância da utilização deste *delay* de escala de tempo, reforçando que, mesmo trabalhando em escala mensal, os dados de precipitação de meses anteriores agregam informações à adaptatividade da RNA. Para trabalhos futuros recomenda-se mais variações dos dados de entrada, com a utilização de outras características meteorológicas, além da precipitação, como evapotranspiração potencial, evaporação, insolação etc.

Corroborando com o que Kaviski (1992) concluiu, a análise de agrupamentos, quando combinada com a técnica de análise discriminante, permite que se tenha um critério objetivo para identificar a região homogênea para a qual pertence um local sem dados observados. No caso deste trabalho, pelo resultado dos grupos formados, verificou-se que, com exceção da estação Terresa Cristina incluída no Grupo 2, a localização geográfica das estações teve grande influência na formação dos grupos. Enquanto o Grupo 1 foi formado por estações inseridas na porção mais a oeste da abrangência geral estudada, o Grupo 3 formou-se por estações inseridas mais ao leste do mesmo espaço, restando ao Grupo 2 as estações mais ao centro.

Como o Paraná se caracteriza pela baixa sazonalidade de chuva e vazão, é plausível considerar que parâmetros fisiográficos tenham tido influência maior no momento da formação dos grupos, principalmente porque foi considerada a variável longitude na análise. Recomenda-se, portanto, que se use análise de agrupamentos para identificar as regiões homogêneas e análise discriminante para estimar a probabilidade de que um local sem dados observados pertença em cada região homogênea.

Uma maneira de tornar ainda mais detalhada e passível de melhor discriminação a amostra de informações das estações fluviométricas analisadas, e de suas respectivas áreas de drenagem, seria utilizando mais parâmetros para sua caracterização, como fez Diniz (2008), que analisou mais de 30 parâmetros físicos e climatológicos, divididos em características das medidas lineares, características da forma, da rede de drenagem, do relevo e da capacidade de escoamento das bacias hidrográficas.

Com relação ao método proposto para regionalização de séries de vazões mensais utilizando RNAs, o mesmo se mostrou eficiente quando desenvolvido para pequenas bacias hidrográficas do Estado do Paraná, isso por conta de os valores obtidos para aferição dos resultados terem sido consideravelmente semelhantes.

Ao comparar as vazões regionalizadas com as observadas nas bacias teste, os resultados foram bastante satisfatórios. No caso da estação Guampará, o coeficiente de Nash-Sutcliffe resultou em 0,82, mesmo valor obtido para o coeficiente de determinação. Além disso, as médias e desvios de ambas as séries foram bem semelhantes, com diferença de 7,5% para as mé-

dias e 9,9% para os desvios padrão da série regionalizada para a observada. Para a estação Madeireira Gavazzoni, os resultados foram mais expressivos ainda, os coeficientes de Nash-Sutcliffe e de determinação tiveram também o mesmo valor, resultando em 0,88, e as diferenças entre as médias e desvios foram de 0,5% e 9,4%, respectivamente.

Ter utilizado o método do inverso do quadrado da distância para a regionalização das vazões obtidas das estações do mesmo grupo para a estação base trouxe bons resultados ao trabalho, pois se verificou que as estatísticas obtidas ao aplicar a ponderação na série de vazões resultaram em valores ligeiramente maiores do que se a série fosse regionalizada baseada simplesmente na RNA de cada uma das estações do grupo, isoladamente.

Além disso, não foram verificados erros sistêmicos de majoração ou minoração das séries, mostrando que as RNAs utilizadas estavam treinadas para toda a amplitude de valores inseridos.

Concluindo, por se tratar de uma técnica objetiva de regionalização hidrológica, o método proposto pode ser implementado em tempo mínimo, com uma utilização de maneira simples, rápida e segura.

Recomenda-se, portanto, para trabalhos futuros, a expansão da aplicação do método de regionalização estudado, ou seja, sua aplicação para mais bacias hidrográficas nas diversas regiões do país (sul, sudeste, centro-oeste, norte e nordeste), de modo a criar um banco de dados nacional com informações de diversos estados para servir de base para avaliação prévia de pequenos potenciais hidrelétricos no país inteiro.

Para melhorar ainda mais a aferição dos resultados das RNAs modeladas, recomenda-se o estudo de novos algoritmos de otimização, como os de Regularização Bayesiana, BFGS Quasi-Newton, da Secante, do Gradiente Descendente com Momento etc., todos disponíveis em plataforma MATLAB. Além disso, novas estruturas e modelos de RNAs podem ser testadas, e não apenas a *Multi Layer Perceptron* com três camadas e variados neurônios na camada intermediária, como neste trabalho.

Além disso, devido às inerentes incertezas de estudos realizados em hidrologia, relacionados principalmente às características das bacias hidrográficas e à imprecisão e escassez de dados hidrológicos, recomenda-se a utilização de análise difusa na regionalização de parâmetros hidrológicos e no trei-

namento e validação das RNAs, pois se trata de um método de multicritérios, para impactos sob incerteza, que pode trazer melhores resultados.

## 7 REFERÊNCIAS BIBLIOGRÁFICAS

- ANMALA, J., ZHANG, B., & GOVINDARAJU, R. S. (2000). Comparison of ANNs and empirical approaches for predicting watershed runoff. *Journal of Water Resources Planning and Management* v. 126. n.3 , 156-166.
- CEHPAR. (1989). *Aproveitamentos Hidrelétricos de Pequeno Porte. Projeto HG-52, Anexo I*. Curitiba.
- CEHPAR. (2006). *Comparação das curvas de permanência obtidas pelo método de desagregação de vazões diárias com as obtidas pelo modelo de regionalização - HG-171*. Curitiba: I Simpósio de Recursos Hídricos do Sul-Sudeste.
- DINIZ, L. S. (2008). *Regionalização de parâmetros de modelos chuva-vazão usando redes neurais*. Porto Alegre: Tese de Doutorado. Programa de Pós Graduação em Recursos Hídricos e Saneamento Ambiental da Universidade Federal do Rio Grande do Sul.
- ELSHORBAGY, A. S., & PANU, U. S. (2000). Performance evaluation of artificial neural networks for runoff prediction. *Journal of Hydrologic Engineering* v.5 n.4, 424-427.
- FARAWAY, J., & CHATFIELD, C. (1998). Time series forecasting with neural networks: a comparative study using the airline data. *Applied Statistics*, 231-250.
- FERREIRA, D. F. (2008). *Estatística Multivariada*. Lavras: UFLA.
- FILL, H. (1987). Informações hidrológicas. In: F. BARTH, C. POMPEU, H. FILL, C. TUCCI, J. KELMAN, & B. BRAGA JÚNIOR, *Modelos para gerenciamento de recursos hídricos*. São Paulo. Coleção ABRH de Recursos Hídricos.
- FORTIN, V. O. (1997). Comments on 'The use of artificial neural networks for the prediction of water quality parameters'. *Water Resources Research* , 2423-2424.
- GOMES, J. (1977). Investigação sobre métodos de extrapolação da curva de descarga. *Simpósio Brasileiro de Recursos Hídricos (1997)*, 509-518. Vitória: ABRH.
- GOMES, L. F., & MONTENEGRO, S. M. (2010). Modelo baseado na técnica de redes neurais artificiais para previsão de vazões na bacia do rio São Francisco, *ABRH v.15 n.1*, 5-15.
- HAGAN, M. T., & B., M. M. (1994). Training feedforward newtorks with the Marquardt Algorithm. *IEEE Transactions on Neural Networks* , 989-993.

HSU, K. L., GUPTA, H., & SOROOSHIAN, S. (1995). Artificial neural network modeling of rainfall-runoff process. *Water Resources Research* v.31 n.10, 2517-2530.

HUGHES, J. M. (1987). Hydrological characteristics and classification of Tasmanian rivers. *Aust. Geog. Studies* v.25, 61-82.

KAVISKI, E. (1992). *Métodos de regionalização de eventos e parâmetros hidrológicos*. Curitiba: Dissertação de mestrado.

KAVISKI, E., KRÜGER, C. M., & ILLICH, I. (1993). Regionalização de vazões médias em pequenas bacias hidrográficas do Estado de Santa Catarina. *Anais do X Simpósio Brasileiro de Recursos Hídricos e I Simpósio de Recursos Hídricos do Cone Sul*.

KRÜGER, C. M., & KAVISKI, E. (1994). Regionalização de Vazões em Pequenas Bacias no Estado do Paraná. Curitiba: Centro de Hidráulica do Paraná.

LIMA, C. H., & FERREIRA FILHO, W. M. (2003). Análise de modelos de redes neurais aplicados ao processo chuva-deflúvio no semi-árido. *XV Simpósio Brasileiro de Recursos Hídricos*, Curitiba.

LOPES, J. E., BRAGA, B. P., & CONEJO, J. G. (1981). Simulação hidrológica: aplicações de um modelo simplificado. *IV Simpósio Brasileiro de Recursos Hídricos*, 42-62. São Paulo.

MACHADO, F. W. (2005). *Modelagem chuva-vazão utilizando redes neurais artificiais*. Curitiba: Dissertação de mestrado.

MACHADO, F. W., MINE, M. R., KAVISKI, E., & FILL, H. D. (2011). Monthly rainfall-runoff modelling using artificial neural networks. *Hydrological Sciences Journal* v. 56 , 249-361.

MAIDMENT, D. (1992). *Handbook of hydrology*. Texas.

MAIER, H. R., & DANDY, G. C. (2000). Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling & Software* v. 15, 101-124.

MATLAB. (2012). *Neural network toolbox documentation*.

MATLAB. (2012). *Statistics toolbox*.

MINGOTI, A. A. (2005). *Análise de Dados Através de Métodos de Estatística Multivariada*. UFMG, Belo Horizonte.

MOSLEY, M. P. (1981). Delimitation of New Zealand hydrologic regions. *Journal of Hydrology* v. 49, 173-192.

NOVAES, L. F., & PRUSKI, F. F. (2007). Avaliação do desempenho de cinco metodologias de regionalização de vazões. *ABRH* v. 536 .

RAJURKAR, M. P., KOTHYARI, U. C., & CHAUBE, U. C. (2002). Artificial neural networks for daily rainfall-runoff modelling. *Hydrological Sciences Journal* v.47, 865-877.

RAJURKAR, M. P., KOTHYARI, U. C., & CHAUBE, U. C. (2004). Modelling of daily rainfall-runoff relationship with artificial neural network. *Journal of Hydrology* v.285, 96-113.

ROWEIS, S. (2003). Levenberg Marquardt Optimization. (disponível em [www.cs.toronto.edu/~roweis/notes/lm.pdf](http://www.cs.toronto.edu/~roweis/notes/lm.pdf)).

SAJIKUMAR, N., & THANDAVESWARA, B. (1999). A non-linear rainfall-runoff model using artificial neural network. *Journal of Hydrology* v. 216 n.1-2, 32-55.

SANTOS, I., FILL, H., SUGAI, M., BUBA, H., KISHI, R., MARONE, E., et al. (2001). *Hidrometria Aplicada*. Curitiba: LACTEC.

SARAIVA, I. (2010). *Regionalização de parâmetros de modelos de simulação hidrológica para emprego em bacias hidrográficas sem monitoramento fluviométrico*. Belo Horizonte. Dissertação de mestrado.

SEFIONE, A. L. (2002). *Estudo comparativo de métodos de extrapolação superior de curvas-chave*. Porto Alegre: Dissertação de mestrado.

SMAKA, C. (2010). *Aplicação de análise multivariada na identificação de fatores que influenciam no custo de um plano de saúde*. Acesso em 08 de Junho de 2013, disponível em <http://www.ppgmne.ufpr.br/arquivos/diss/229.pdf>

SPERB, R. M., MODRO, N. R., MARCONDES, P., M., FOES, P. A., BARCIA, R. M., PACHECO, R. C. (1999). Prevent: Protótipo de um sistema de previsão de enchentes baseado em redes neurais. XIII *SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS*.

STARK, P. (1979). *Introdução aos métodos numéricos*. Rio de Janeiro.

TASKER, G. D. (1982). Comparing methods of hydrologic regionalization. *Water Resources Bulletin*, 965-970.

TASKER, G. D. (1982). Simplified testing of hydrologic regression regions. *J. Hydraul. Div.*, 128-1222.

TOKAR, A. S., & JOHNSON, P. A. (1999). Rainfall-runoff modeling using artificial neural networks. *Journal of Hydrologic Engineering* v.4 n.3, 232-239.

TOKAR, A. S., & MARKUS, M. (2000). Precipitation – runoff modelling using artificial neural networks and conceptual models. *Journal of Hydrologic Engineering* v.5 n.2, 156-161.

TUCCI, C. E. (1998). *Modelos Hidrológicos*. Porto Alegre.

TUCCI, C. E. (2002). *Regionalização de vazões*. Porto Alegre.



TUCCI, C. E., & SILVEIRA, R. L. (1985). *Análise de consistência de dados fluviométricos*. Porto Alegre.

VALENÇA, M. J. (1999). *Análise e projeto de redes neurais com abordagem construtiva para modelagem de sistemas complexos*. Recife: Tese de doutorado, UFPE/CCEN/Ciência da Computação.

VENDRUSCOLO, J. (2005). *Regionalização dos parâmetros de um modelo chuva-vazão para estimativa de vazões em bacias sem monitoramento sistemático*. Santa Maria: Dissertação de mestrado.

WILKS, D. S. (2011). *Statistical methods in the atmospheric sciences*. Ithaca: ELSEVIER.

## APÊNDICE 1 – RESULTADOS DA MODELAGEM CHUVA-VAZÃO

Os gráficos a seguir mostram a comparação entre os hidrogramas observados no ponto de localização das estações fluviométricas consideradas neste estudo e os hidrogramas simulados utilizando as Redes Neurais Artificiais como modelo chuva-vazão.

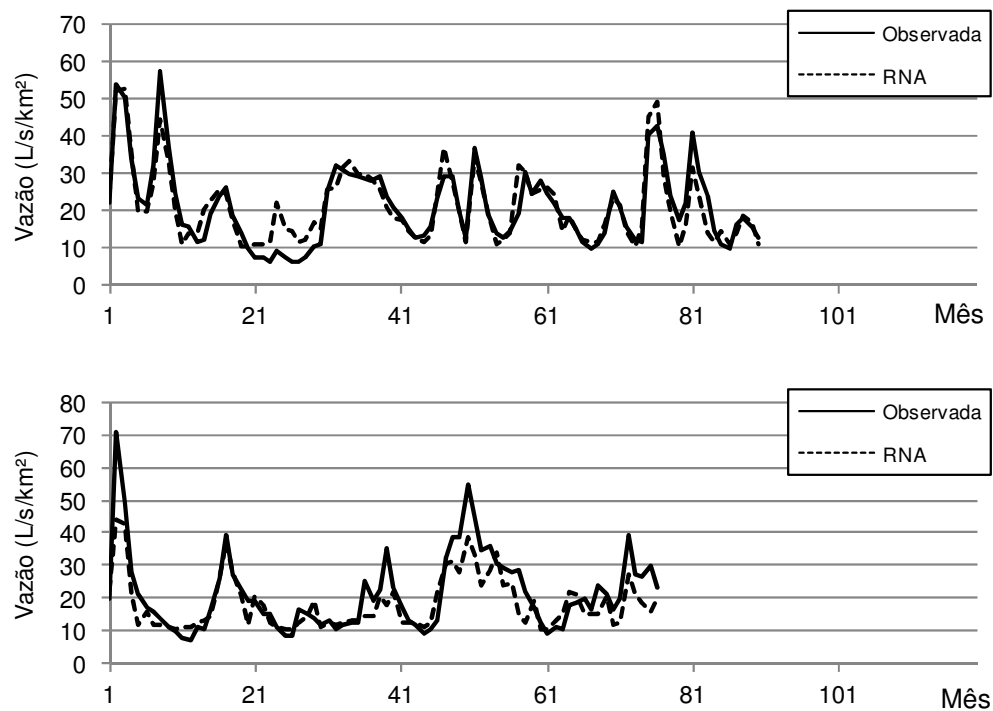


FIGURA 1.1: CALIBRAÇÃO E VALIDAÇÃO DAS RNAs PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO PONTE DO GOIO BANG.

FONTE: O AUTOR (2014).

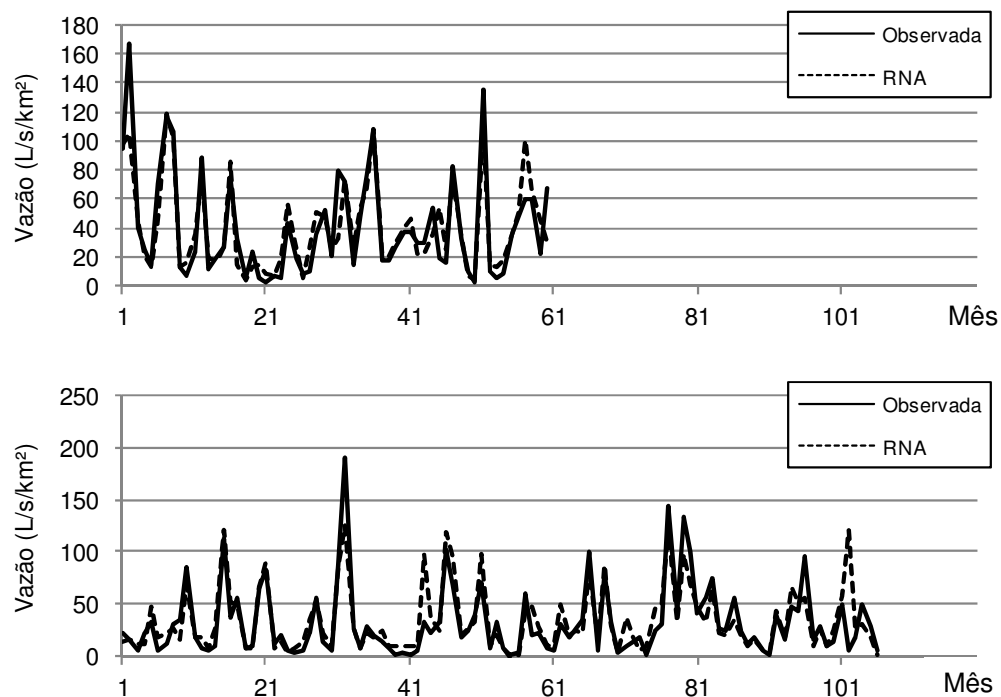


FIGURA 1.2: CALIBRAÇÃO E VALIDAÇÃO DAS RNAs PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO PONTE LEÔNICO PRIMO.

FONTE: O AUTOR (2014).

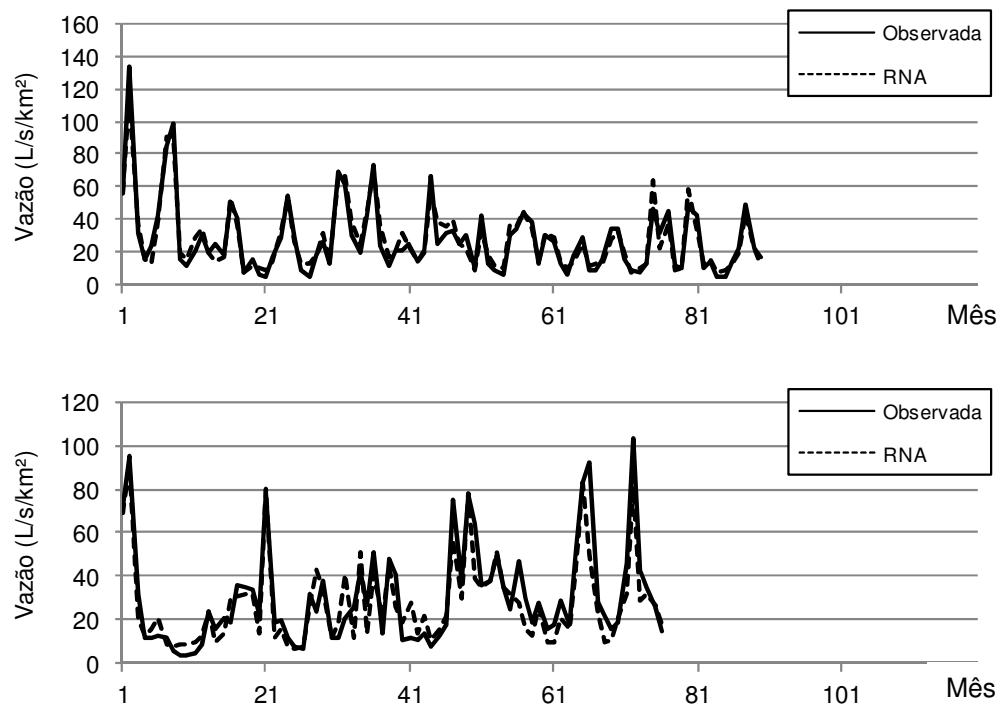


FIGURA 1.3: CALIBRAÇÃO E VALIDAÇÃO DAS RNAs PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO ETA - GUARAPUAVA.

FONTE: O AUTOR (2014).

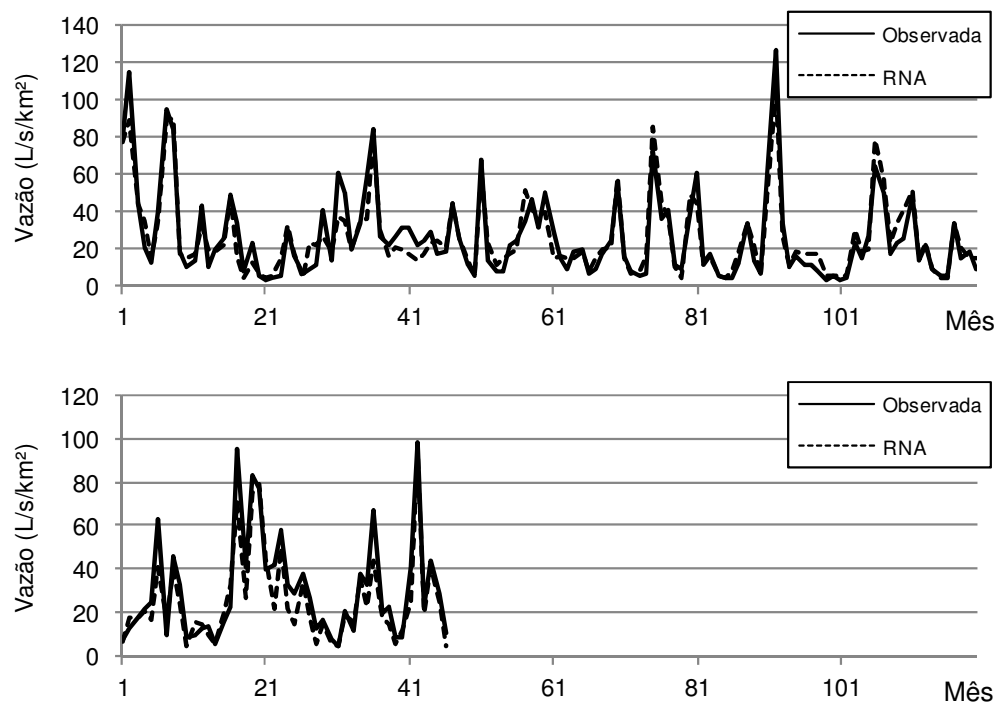


FIGURA 1.4: CALIBRAÇÃO E VALIDAÇÃO DAS RNAs PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO BALSA DO CANTÚ.

FONTE: O AUTOR (2014).

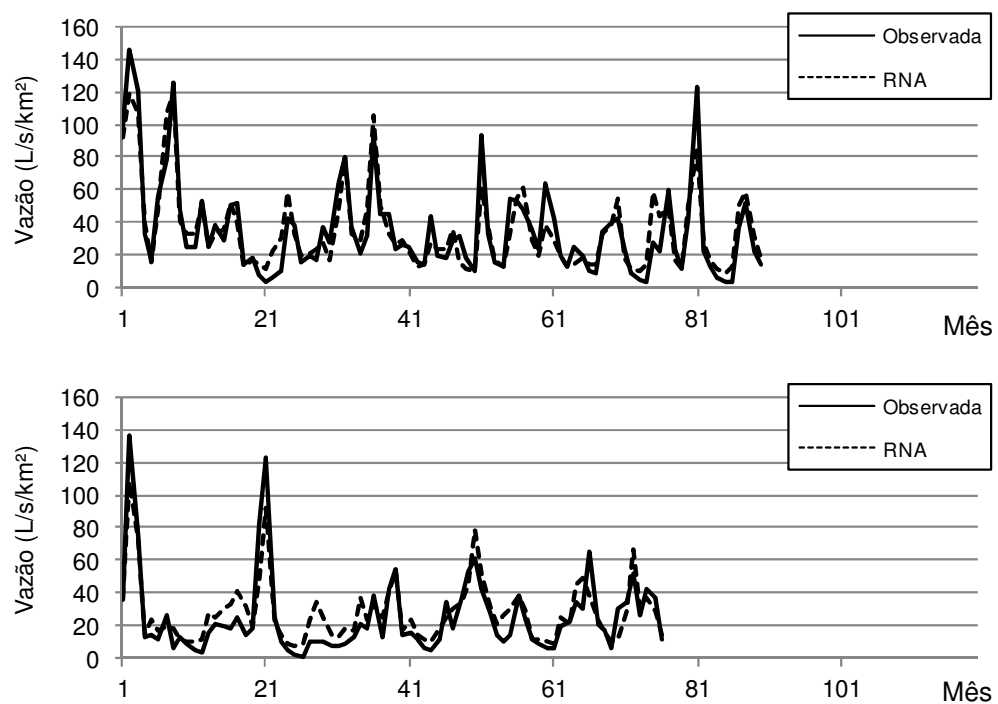


FIGURA 1.5: CALIBRAÇÃO E VALIDAÇÃO DAS RNAs PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO TORMENTA FAZENDA TRENTO.

FONTE: O AUTOR (2014).

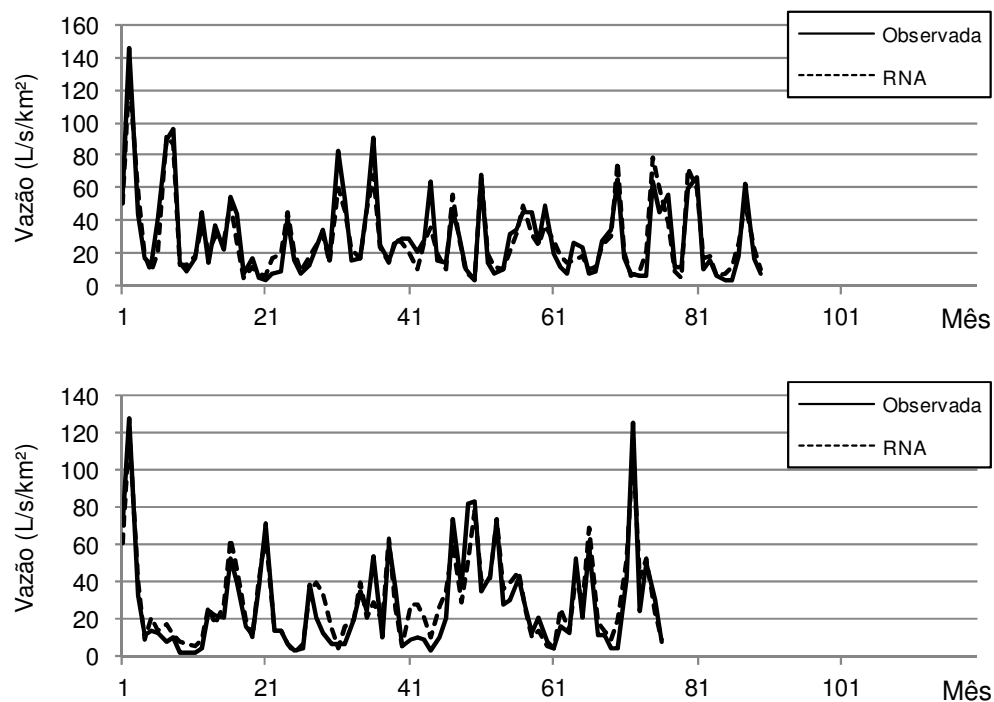


FIGURA 1.6: CALIBRAÇÃO E VALIDAÇÃO DAS RNAs PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO PORTO CARRIEL.

FONTE: O AUTOR (2014).

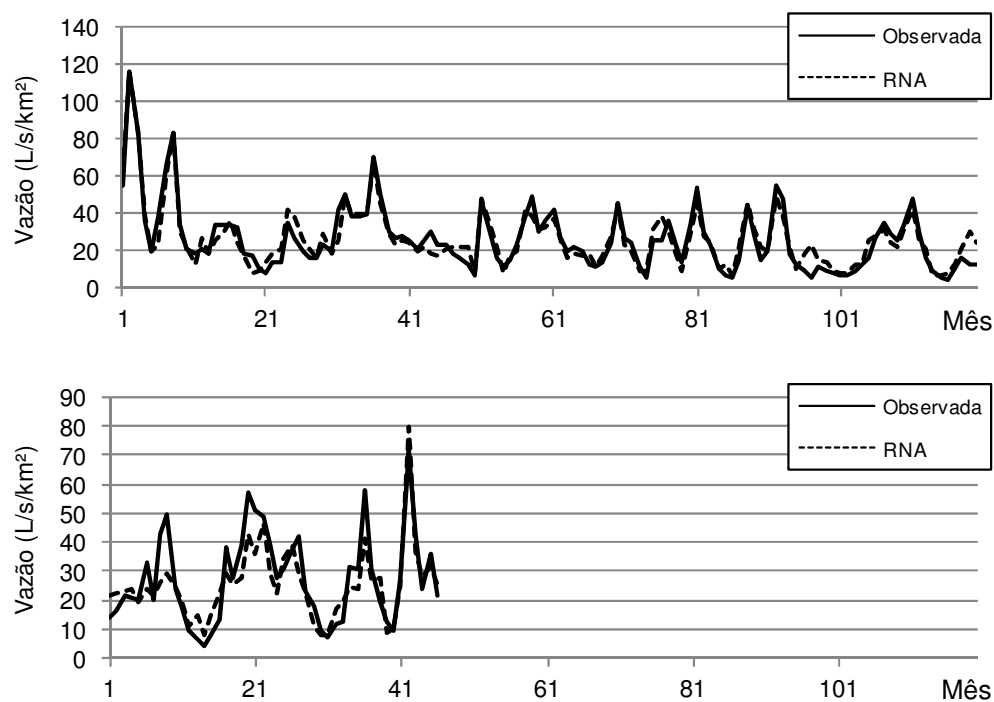


FIGURA 1.7: CALIBRAÇÃO E VALIDAÇÃO DAS RNAs PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO SALTO SAPUCAÍ.

FONTE: O AUTOR (2014).

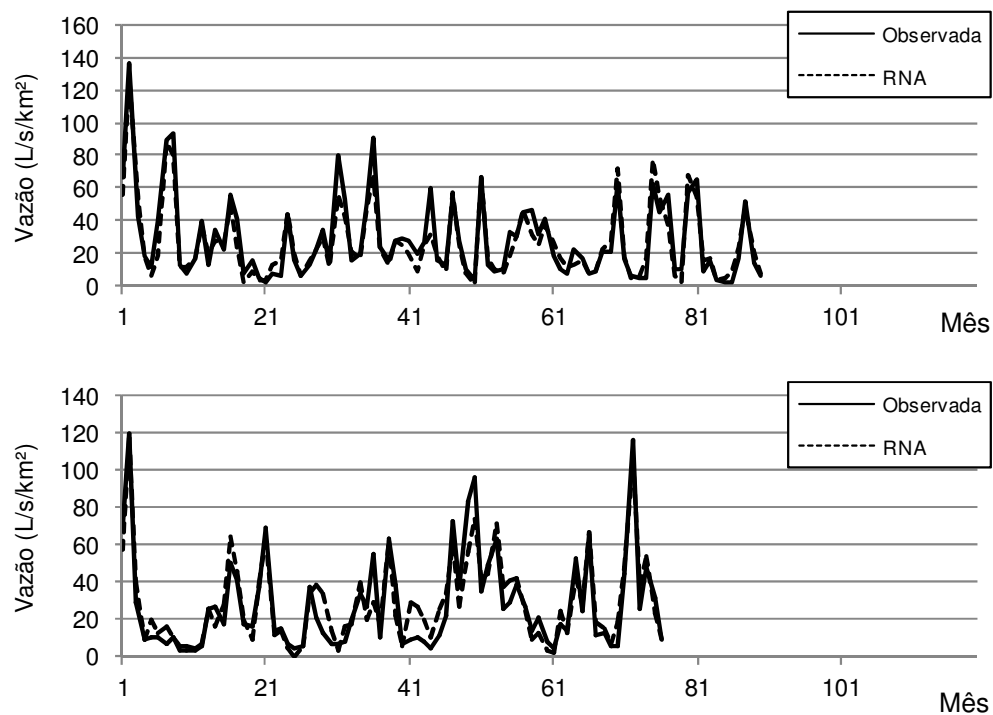


FIGURA 1.8: CALIBRAÇÃO E VALIDAÇÃO DAS RNAs PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO PORTO GUARANI.

FONTE: O AUTOR (2014).

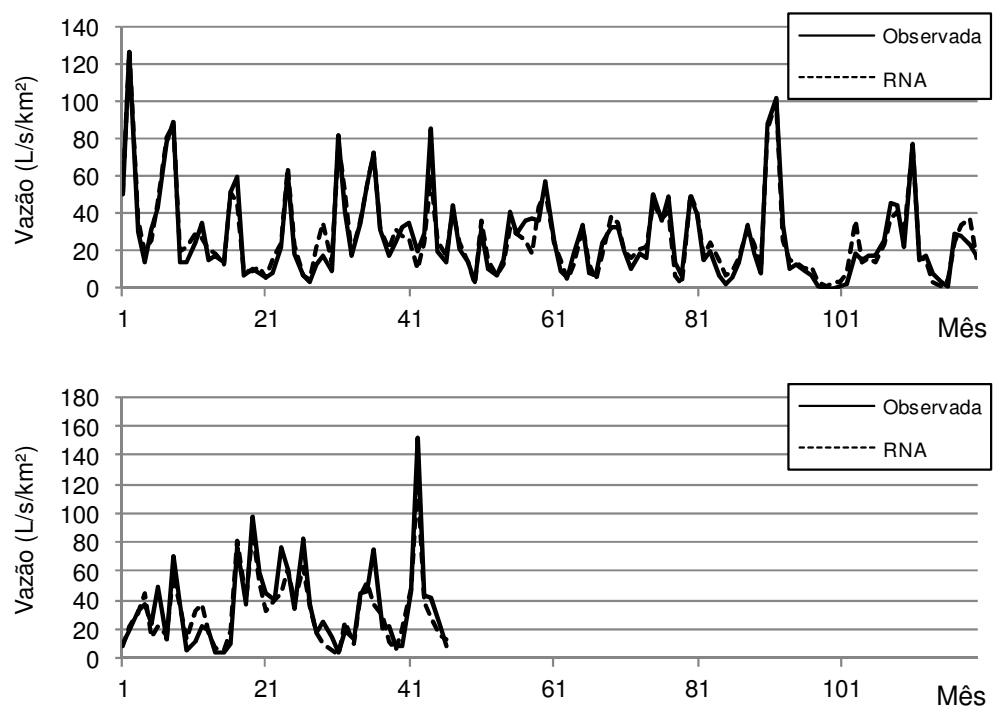


FIGURA 1.9: CALIBRAÇÃO E VALIDAÇÃO DAS RNAs PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO RIO DOS PATOS.

FONTE: O AUTOR (2014).

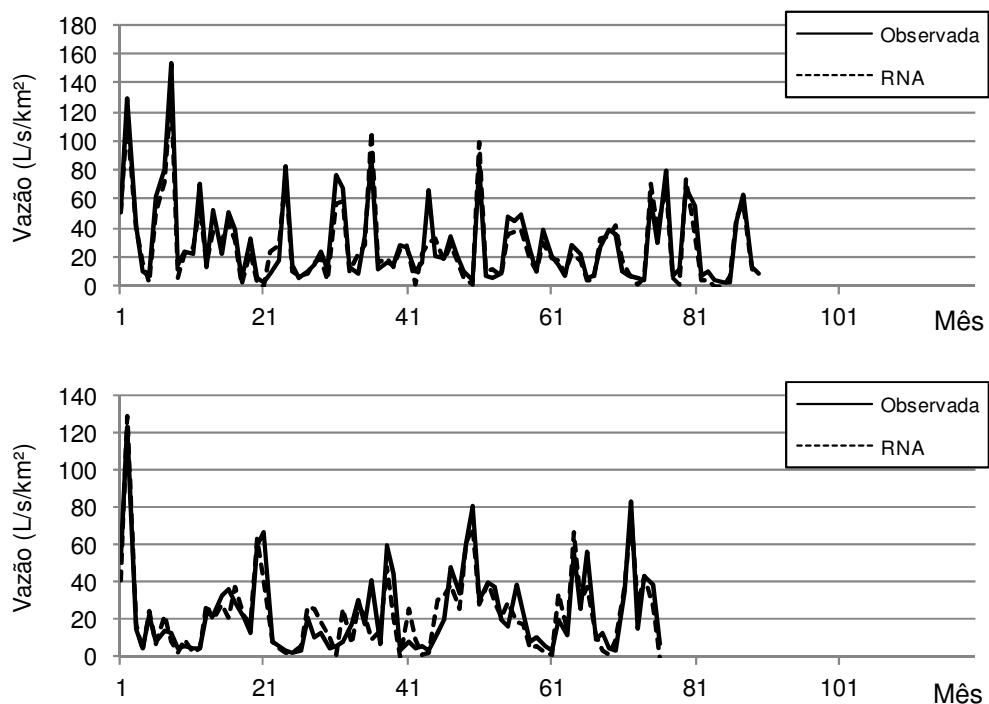


FIGURA 1.10: CALIBRAÇÃO E VALIDAÇÃO DAS RNAs PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO PORTO SANTO ANTÔNIO.

FONTE: O AUTOR (2014).

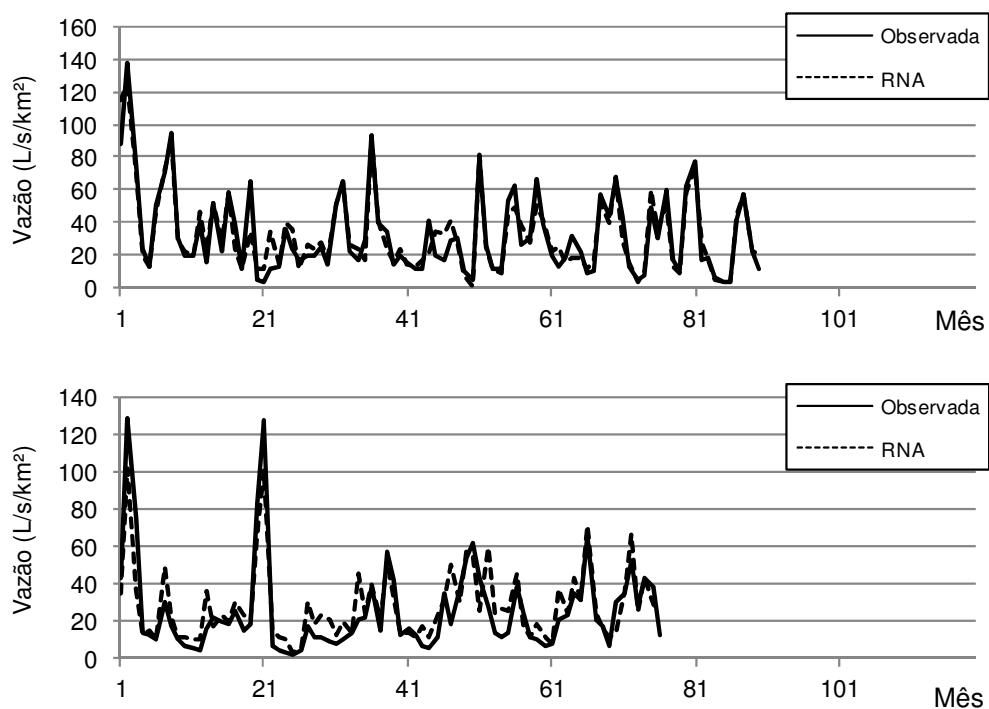


FIGURA 1.11: CALIBRAÇÃO E VALIDAÇÃO DAS RNAs PARA A SÉRIE HIDROLÓGICA DA ESTAÇÃO SÃO SEBASTIÃO.

FONTE: O AUTOR (2014).